THE URBAN INSTITUTE

2100 M Street, N.W.
Washington, D.C. 20037

Project Report

3075-4 February 1983

Alternative Methods of Developing a Relative Value Scale of Physicians' Services: Year 1 Report

This report was prepared by Jack Hadley, David Juba, Katherine Swartz, Judith Wagner, and Robert Berenson. Opinions expressed are those of the authors and do not necessarily represent the views of The Urban Institute or its sponsors. The authors thank Ellen Pisarski and Cathy Carlson for their excellent research assistance.

Contract Information

Contract no.	HCFA-500-81-0053
Type and amount of award	Competitive, $287,550
Awarding agency	ORDS, HCFA, HHS Contract Branch, Area D-2, GOB 6325 Security Blvd. Baltimore, MD 21207
Project Officer	James R. Cantwell
Project Director	Jack Hadley

THE URBAN INSTITUTE
2100 M Street, N.W., Washington, D. C. 20037

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

CHAPTER 1

GOALS, PURPOSES AND HISTORY OF RELATIVE VALUE SCALES

The Health Care Financing Administration is in the process of reevaluating how it pays for medical care. It is considering prospective payment methods for both hospitals and skilled nursing facilities. It is granting states considerably more authority to implement so-called alternative reimbursement systems that include Medicare and Medicaid. Payment of physicians' services has been evolving from an essentially open-ended, inherently inflationary system, i.e., the UCR method, to a _de facto_ fee schedule because of the Economic Index. As more and more physicians' Medicare-computed reasonable charges for specific services exceed the ceilings imposed by the Economic Index, then the ceilings become, in effect, fixed fees paid to all physicians whose reasonable charges exceed the ceiling.

The evolution of a _de facto_ fee schedule through the Economic Index will certainly help HCFA attain one of its objectives, limiting its payments for physicians' services. However, this fee schedule is unlikely to reflect very accurately the _relative_ values of different physicians' services (Paringer, 1981b). Since value should incorporate both the costs of producing services and patients' (and insurers') preferences, all parties are likely to be unhappy with a _de facto_ fee schedule. Physicians will be unhappy because Medicare's payments may bear little relation to the costs of providing care. This should further erode physicians' willingness to accept assignment of benefits (Paringer, 1981a). Beneficiaries will be unhappy because their cost sharing, especially for charges in excess of Medicare payments, will be both greater and haphazardly related to the types of services they receive and their perceptions of the values of those services. HCFA will be unhappy for two reasons: it will have to deal with unhappy physicians and beneficiaries and it will have little ability to influence relative fees.*

The purpose of this project is to explore and evaluate alternative methods of constructing relative value scales of physicians' services. A relative value scale is not a fee schedule, though it is an important intermediate step in developing a fee schedule. Given a procedure nomenclature that identifies and defines all the different services a physician can provide, a relative value scale is a cardinal ranking of those services. This means two things. First, each service's numerical assignment orders that service in relation to all other services. Second, the difference between any two services' numerical assignment measures the difference in some concept of value. In other words, a relative value scale enables one to say something like, "One service is worth ten times another, or is ten times more valuable, or is equivalent to ten units of another."

Although this seems reasonably straight-forward, defining "value" is a difficult task. Ideally, a system of relative values should reflect patients' preferences as well as providers' costs. Otherwise, patients may have incentives to use too much or too little of a type of service, or physicians may have incentives to provide too much or too little.

Values should be consistent with physicians' expenses so that physicians will be willing to provide patients with medically appropriate services. This means that they should be related to the costs of physicians' own time, to the costs of other employees' time, to the costs of equipment and facilities, and

---

*The latter is important if physicians' decisions are influenced by relative fees. Although there is no good evidence regarding the impact on specific services, several studies have shown that other key decisions are influenced by financial incentives (Sloan, Cromwell, and Mitchel, 1978; Hadley, 1979; Paringer, 1981; Hadley, 1982).

to the costs of supplies. For some services, the amount of a physician's time used in producing a service will provide a reasonable approximation of costs; for other services, it will not.

It is difficult to separate physicians' costs from patients' preferences. To begin with, the value, or opportunity cost, of a physician's own time depends in part on how much patients and insurors are willing to pay. In areas in which fees are high, physicians will tend to regard the value of their time as being high as well. Secondly, it may not always be desirable to have services produced in the least expensive way because the costs of some services are sensitive to how many are provided. If, for example, a specialized skill or piece of equipment is used in providing a service, the average cost of the service will be lowest when volume is fairly high. Patients, however, may be willing to pay the higher costs of small scale provision to ensure that a service is available locally. This means that there are likely to be variations in the costs of providing services which will not be attributable to variations in the costs of physicians' time, employees' time, supplies, or facilities and equipment.

There is another reason why ideal relative values should reflect more than just the costs of providing services as cheaply as possible: many of the costs associated with treatment are not borne by physicians. These external costs include such things as hospitalization, time off from work or school, and patients' time. Ideally, fees based on relative values should provide incentives for physicians to offer services which limit these external costs, not just their own expenses associated with providing a service.

Thus, relative values should reflect not only the costs of efficient production but also the preferences and costs faced by patients and society. As a practical matter, this may be an unrealistic objective. A more reasonable goal might be an RVS that avoids giving physicians very wrong signals about relative values.

Once relative values of procedures are determined, the construction of a fee schedule is conceptually straightforward. Assigning relative value units a monetary value converts the relative value scale into a fee schedule. Each procedure on the scale would have a monetary amount as well as a "pure" number associated with it. If all relative value units receive the same monetary value, then the schedule of relative fees would be identical to the schedule of relative values. Conversely, if some procedures' relative value units are assigned a greater monetary value than others', then relative fees will not be identical to relative values.

The absolute magnitude of the monetary conversion factor obviously determines the absolute level of fees, e.g., a conversion factor of $10.00 per relative value unit will result in fees ten percent higher than those based on $9.09 per relative value unit. Since it is relative and absolute fees that ultimately influence physicians' and patients' behavior, determining the appropriate financial conversion factor(s) is a key policy decision.

The following chapters of this report describe and evaluate five alternative methods of constructing a relative value schedule: charge-based, physicians' time, statistical cost functions, micro-costing, and consensus development/social preferences. To the extent possible, each evaluation covers the same basic criteria: data requirements; technical expertise requirements; ability to value certain "problem" types of procedures, e.g., new procedures, or procedures undergoing rapid technical change; potential for future updating; feasibility of developing specialty or location-specific scales; feasibility of implementation at the carrier level; and potential for capturing costs and benefits that occur outside the direct patient-physician encounter. In order to set the context for these chapters, the remainder of this chapter presents a brief history of relative value scales and a more detailed discussion of how relative value scales relate to and can potentially influence physicians' fees.

A. HISTORY OF RELATIVE VALUE SCALES

Despite the ease of constructing a fee schedule from a RVS, the two are conceptually distinct. While medical fee schedules were employed in colonial America, becoming widespread during the nineteenth century (Showstack et al., 1979), RVSs are a comparatively recent phenomenon. Impetus for the latter seems to have come from the rapid proliferation of third-party coverage of medical services over the past three decades. Up to the end of World War II, most physicians were reimbursed directly by their patients for the performance of medical services. Such a billing arrangement did not require consistency of prices or charges across physicians or patients. The conventional belief that some physicians engaged in cross-subsidization during this period--charging patients according to ability-to-pay--is a manifestation of these inconsistencies.

Also during this post-war period, overt price competition within the physicians' services market was actively discouraged by medical societies. Consequently, for some physicians and procedures, fees may not have been congruent with marginal production costs. Faced with possibly distorted market price structures, third-party payors could not rely solely upon individual physician's billed charges to determine adequate compensation for the services provided.

With the expansion of third-party coverage, then, physicians and insurers felt increasing pressures to rationalize reimbursement methods and establish consistent and fair levels of payment. Modern attempts to base such schemes upon RVSs go back at least as far as the early 1940s. At that time, the Casualty Actuarial Society developed RVSs at the request of commercial insurers (Showstack et al., 1979). Subsequent RVSs were published by seventeen state medical societies and six national specialty societies (Eisenberg, 1980).

By far the best known RVS was introduced in 1956 by the California Medical Association. Now known as the California Relative Value Studies (CRVS), editions were published in 1956, 1957, 1960, 1969, and 1974. Early versions of the CRVS were based on surveys of how much California physicians "generally charged" for different services. The 1969 and 1974 editions based most relative values on median fee data derived from the files of third-party payors operating within the state.

Each subsequent edition of the CRVS increased the number of procedures listed; three factors contributed to this trend. One was the introduction of new services. A second was the splitting or "unpackaging" of services into component parts. The third was the increasing use of modifiers to distinguish among procedures of different complexity and length.

Other RVSs were constructed by professional societies like the American College of Obstetricians and Gynecologists, and the American Academy of Orthopedic Surgeons. Blue Shield plans and commercial insurance companies, too, developed RVSs, although many seem to have a close resemblance to the CRVS (Hsiao and Stason, 1979).

During the latter half of the past decade, anti-trust actions initiated by several governmental agencies stopped the development and publication of professional societies' scales including the CRVS series. These same agencies, however, do not seem to view insurers' RVSs with the same suspicions and have not sought to restrict third-party payors from employing their scales in administering medical insurance (Showstack et al., 1979).

Still other governmental agencies, especially HCFA because of its responsibility for the administration of Medicare and Medicaid, take a more positive approach in the matter of fee schedules and relative value scales. These agencies' interest in the scales is dictated by the large public budgets which the health care sector consumes on a regular and growing basis. While Medicare payment is based primarily on a "usual and customary" (UCR) system reflecting Medicare relative fees, maximum fee levels have always been a part of this system. Therefore, proper relative fees have an importance to Medicare even under UCR reimbursement. In addition, many states use RVSs in determining payment to physicians for services provided to Medicaid patients (Holahan, Gornick and Nichols, 1981).

B. RVS AND RELATIVE PRICES

There are three dimensions to the relationship between an RVS and a system of relative prices for physicians' services. These are:

- The definition of procedures through the coding and terminology employed.
- The assignment of relative values to procedures which have been defined by the coding and terminology system.
- The impact of the resulting RVS on physicians' relative prices.

Each dimension of this relationship may affect not only fee-for-service reimbursement of physicians, but other parts of the health care sector as well.

Coding and Terminology

How a coding and terminology system defines procedures has implications for patients, physicians, and third-party payors. For example, coding and terminology affect physicians' incentives. Physicians may use detailed coding systems to increase revenues, even though fee screens or price controls are designed to limit such increases (Sobaski, 1975; Holahan and Scanlon, 1978). It has also been suggested that some coding and terminology systems may permit, or even encourage, physicians to increase revenues by unpackaging--billing for several individual services instead of one inclusive service (Kirchner, 1979). These manipulations of the coding and terminology system represent only one of several responses to attempts to control program fees; depending on the structure of the program, other responses may increase patients' costs or make access more difficult (Lee and Hadley, 1979).

It is possible that cost savings might result if codes were defined so that unpackaging were not possible. Creating aggregate codes, i.e., packaging, reduces physicians' incentives to provide additional services; the implications of such efforts are not unambiguous, however. For example, limited evidence suggests that per-case reimbursement for hospitalized patients (one form of packaging) may lead to reduced lengths of stay (Markel, 1980). While this may be a desirable result in some instances, in others it may not. For example, per-case reimbursement of nursing home patients could significantly reduce the number of visits by physicians, an undesirable outcome in some cases.

Coding and terminology can affect both patients' and physicians' incentives. For example, some forms of surgery can be performed on either an inpatient or outpatient basis. If the coding system differentiates between the two and if, as is common, insurance coverage is more complete for inpatient than outpatient care, an insured individual may find it advantageous to have the procedure performed as an inpatient, even though total costs may be higher.

Assigning Relative Values to Procedures

Once procedures have been defined, the second step in constructing an RVS is the actual assignment of a relative value to each procedure. The California Relative Value Studies, for instance, computed most procedures' relative values by comparing median charges to the median charge of some numeraire or baseline procedure. Given the original purpose of the CRVS, this was a perfectly reasonable approach to follow. If, however, the purpose of the RVS is to influence or set relative fees, then simply basing relative values on relative fees may not be desirable in all cases.

Alternative approaches vary in terms of underlying objectives of the RVS and method of construction. For example, one set of approaches focuses primarily on the resource costs of providing physicians' services. However, within this class one could further distinguish between relative values based on the full costs to society, the full costs to the individual physicians, or simply the time costs to the physician. The last, of course, is a subset of physicians' full costs. The first differs from the second in that the full social cost of treating a case should include the cost of drugs, hospital stays, and ancillary services. Use of these other resources is generally controlled by physicians, although the physician is not responsible for paying their costs. Given that the relative importance of non-physician costs and of physician time vary from procedure to procedure, it should be clear that relative values based on these alternative measures of costs will also vary.

Whatever concept of resource cost is chosen, a number of options for constructing an RVS exist. In particular, cost data can be collected by surveys or interviews, by examining physicians' accounts, or by doing time and motion studies or task analyses to measure the physician inputs employed in providing each type of procedure. (Inputs would then be priced at going rates to determine costs.) Once the basic data are obtained, then various methods such as statistical cost functions, micro-costing, or Delphi methods could be used to determine efficient production costs.

A second major set of approaches focuses on a broader definition of value. By their nature the resource cost methods tend to ignore patients' costs, such as travel and waiting time, and broader social preferences regarding the mix and distribution of services. Although social preferences are difficult, if not impossible, to measure precisely, statements from the Congress, DHSS, the Institute of Medicine, and other authoritative bodies suggest that the current mix of physicians' services deviates from the social ideal. In particular, it is thought that there is too much surgery, not enough primary care, too much emphasis on procedure-oriented services, too ready acceptance of new and often very expensive technologies, and too few services in rural and inner-city areas. Modifying relative prices in order to take account of patients' costs and/or social preferences represents another approach to constructing an RVS.

As with resource cost approaches, there are a number of alternative methods of constructing an RVS. The Delphi technique could be used to identify social preferences and/or assign relative values to procedures. Surveys of patients' travel and waiting costs could be added to estimates of full resource costs. One could use "willingness-to-pay" studies or "revealed-preference" studies of informed patients in order to estimate the values of various medical procedures. Or one could simply alter relative prices so as to reflect perceived social preferences.

The purpose of the foregoing was to present a brief overview of the possible alternative methods of constructing RVSs, which is the focus of this report. In general, assigning relative values to procedures depends upon two issues: the ultimate goal or purpose of the RVSs and the methods available (and used) of obtaining and evaluating relevant data. Methods and goals frequently overlap but, at the same time, can diverge significantly in terms of costs and implications. It should be emphasized, however, that there is no reason why multiple methods should not be used to construct an RVS.

The Effect of an RVS on Prices

An RVS can affect physicians' fees directly or indirectly. Direct effects ensue if third-party payors use an RVS to set physicians' fees or if physicians use an RVS in setting fees. A number of states use an RVS plus a conversion factor to calculate fee schedules for Medicaid patients (Gornick and Nichols, 1981). Such fee schedules are usually well below schedules of billed prices, so participation by physicians in these fixed-fee programs is a problem (Hadley, 1978; Lee and Hadley, 1979; Paringer, 1980).

The extent to which physicians use an RVS to set prices is difficult to establish. A concern has been that use of an RVS to set prices smacks of price-fixing. In fact, the RVS multiplier determines absolute prices, even if the RVS is followed explicitly. So the mere use of an RVS does not appear sufficient to establish the existence of price fixing per se, since individual physicians may use different multipliers. Through its effects on third-party payors' fee screens, however, an RVS may lead to a misalignment of relative prices, provision of inappropriate services, and higher costs.

Indirect effects of an RVS result if the system distorts patients' financial incentives or if the RVS determines third-party payors' reimbursement screens. Either would have an impact on relative prices (Lee and Hadley, 1980). Both types of indirect effects are complex, as factors besides the RVS are likely to play a role.

Despite a number of potential distortions, it is likely that relative prices will reflect relative social values for at least some procedures. Both patients' and physicians' reactions tend to produce a systematic relationship between costs and prices: aside from cross-subsidies, physicians should strive to equate marginal revenues with marginal costs for individual services. Most patients face some out-of-pocket expense, so patients are not completely unresponsive to prices which are out of line (Lee and Hadley, 1979). Whether they be utility maximizers or satisficers, physicians will be better off if they price their services according to this rule (Lee and Hadley, 1979).

This does not mean that the ensuing system of relative prices will necessarily be desirable. If insurance has kept some prices too high, this rule means that physicians will have incentives to produce more of these services. Until the costs of providing these services have been pushed up, physicians will have an incentive to supply more. One way costs are pushed up is by provision of the service by physicians who cannot provide it inexpensively (perhaps due to low volume or limited expertise). Thus, despite the exceptions, relative prices may offer a sound guide for relative values for some procedures. Moreover, prices are routinely collected by third-party payors. Consequently, relative prices are likely to have a role in construction of an RVS scale.

C. CONCLUDING OBSERVATIONS

The ultimate utility of any RVS to the federal government--and particularly of any scale developed as part of this project--will hinge on the scale's usefulness vis-a-vis developing reasonable reimbursement levels for physicians' services. Ideally, a RVS should have the capacity to serve as the foundation for a fee schedule (or maximum allowable reimbursement levels) which would adequately compensate physicians for providing services efficiently and, at the same time, minimize the possibility that physicians would financially benefit by performing one rather than another equally valued procedure.

Besides adequately compensating physicians for their production costs, a RVS-based reimbursement scheme should have built-in incentives encouraging physicians to provide the style and volume of care preferred by their patients and society-at-large. For instance, an ideal price system might encourage physicians to provide a style of care which minimized such patient-borne external costs as school or work days lost. Additionally, patients may wish that specific medical care be available "on-demand" from their physician; if so, optimal volumes and prices of care may not be at levels corresponding to minimum points on physicians' production cost curves. Therefore, an appropriate relative value scale need not necessarily encourage physicians to produce the volume of care associated with minimum unit cost levels for all procedures. Rather, the scale should encourage them to produce socially optimal volumes of care as efficiently as possible.

As an ideal physician reimbursement scheme would be responsive to the preferences and costs borne by patients, physicians, and society-at-large, it follows that an ideal RVS should also reflect these preferences and costs. In perfect markets, prices would be set at levels equilibrating the volumes of services demanded by consumers with volumes of services efficiently produced by suppliers. In such perfect markets, relative prices themselves would serve as a relative value scale. Medical care markets are, however, imperfect for a host of reasons, ranging from the incompleteness of information required for effective consumer participation--with the resulting consumer dependency upon physician-agents to make medical market place decisions--to the pervasiveness of health insurance, which can distort physicians' and patients' perceptions of the true costs of the medical care purchased. Nevertheless, as was argued earlier, it is still likely that relative prices will reflect social values for at least some procedures. In addition, other advantages of charge-based RVSs (to be discussed in a subsequent chapter) are sufficiently attractive to warrant their development.

RVSs exist primarily for the convenience of third-party payors, both public and private. To process large numbers of bills efficiently, a coding and terminology system is needed. To screen bills for reasonableness, something similar to an RVS is needed.

Existing RVSs, however, have not been developed with an explicit concern for efficient resource allocation. Neither the underlying coding and terminology nor the relative value estimates themselves were designed to give patients and providers incentives to use resources economically. Almost by default, third-party payors have accepted RVSs devised by provider associations. It is not clear how closely these RVSs correlate with an optimal RVS.

The goal of this project is to explore alternative methods of constructing RVSs. These RVSs will be compared with one another, and alternative construction methods will be gauged against an explicit set of evaluation criteria. An RVS developed in this manner could serve not only as a screen for third-party payors but also as a stimulus for efficient use and provision of medical services.

CHAPTER 2

CHARGE-BASED METHODS

by

David A. Juba

A. DEVELOPMENT AND EVALUATION OF CHARGE-BASED RV SCALES

Properties of RV Scales

The first task in constructing charge-based relative value scales is to obtain data on physicians' charges for the medical procedures of interest. Undoubtedly, there will exist at least some variation in procedure-specific charges across physicians and over time. Therefore, strategies must be developed for identifying "representative" values of charges for each procedure, a task undertaken in the Methodology section of this chapter. Suffice it to say that reasonable representative charges may be obtained from particular points on the distributions of charges for each medical procedure.

Suppose that one elects to use median--50th percentile point--values as representative charges for each medical procedure of interest here. These medians, themselves, could serve as measures of the relative value of the procedures. At least two problems are associated with RVSs developed from unadjusted charge data, however.

First, scale values would be sensitive to the source of the charge data employed in construction, especially the year in which the data were collected. The median charge for a particular procedure in 1975 is likely to differ substantially from its corresponding value in 1980 for reasons which include pure price inflation. As one would like a RVS to be independent of such pure price inflation effects, some adjustment must be made to raw charge data if they are to serve as the basis for the scales.

A second problem with using points on the distributions of raw charges as the basis for RVSs is that scale values may vary with the particular point on the distribution selected as "representative." For example, for a particular medical procedure the median (50th percentile) and 75th percentile point values on the charge distribution will almost certainly differ. Hence, RVSs based upon unadjusted median charges will almost certainly differ from scales developed from other points on the distribution.

It is possible to ameliorate both problems and still employ charge data as the basis for a relative value scale by expressing representative charges relative to the charge for some base-line or "numeraire" procedure. Defining the representative charge for the numeraire as \$N and the representative charge for any procedure as \$P, the RVS value for that procedure is

$$RVS_p = \$P/\$N.$$

As the relative values are constructed from the ratios of representative charges, they will remain invariant over time, provided the rate of change in representative charges is constant across all procedures.* Therefore, expressing representative charges per procedure in terms of numeraire-equivalents reduces the sensitivity of the RVS to variation induced by intertemporal shifts in charges.

Similarly, expressing charges relative to a numeraire reduces the likely sensitivity of the RVSs to variations in the selection of representative values from distributions of charges for the medical procedures under study. In fact, if the ratio of the ith and jth percentile points is a constant--or

---

*Define the constant rate of annual price change for all procedures to be a. Also, define base-year representative charges for the numeraire and any selected procedure to be \$N and \$P respectively. At some time t year hence, new representative charges will be \$N·(at) and \$P·(at). Consequently, RV scale values for the Pth procedure in the base year will be \$P/\$N, and, t years in the future will have the equivalent value--\$P (at)/\$N (at) = \$P/\$N.

nearly so--across all distributions of procedures' charges, then the RVS may prove to be nearly invariant with respect to the point on the distribution selected to be representative.*

Ultimately, the similarity across alternative RVSs is an empirical question and one which is explored in this report. The point to be emphasized here is that expressing representative charges in numeraire-equivalent units increases the likelihood of stability of the values assigned to particular procedures across alternative relative value scales.

As should be evident from the specification of charge-based RVSs, the selection of a particular procedure to serve as numeraire will have no effect upon the ordinal ranking of procedures on the RVSs. It is also easy to demonstrate that the cardinal values of procedures on alternative charge-based

---

*Define $\$N_i$ and $\$P_i$ as the ith percentile points on the charge distributions for the numeraire and any other relevant procedure. Suppose that for any other percentile point j on these distributions

$$\$N_j = k_{ij}(\$N_i) \text{ and } \$P_j = k_{ij}(\$P_i)$$

If so, the RV scales generated will be invariant with respect to the distributional point selected as representative, since

$$\$P_j/\$N_j = k_{ij}(\$P_i)/k_{ij}(\$N_i) = \$P_i/\$N_i$$

Such a constant relationship between the ith and jth percentile points on charge distributions is unlikely to hold for all procedures of interest. If, however, the ratio $\$P_j/\$P_i$ is nearly constant across many procedures, then alternatively constructed RVSs (while not invariant) will be relatively insensitive to the selection of the representative point in the distribution of charges.

RVSs (constructed from alternative numeraire procedures) will be preserved up to a multiplicative constant.*

One implication of the foregoing is that the selection of base-line procedure can be made arbitrarily. Therefore, for convenience, we shall choose a frequently performed procedure to serve as numeraire for the various charge-based scales. It is straightforward to demonstrate that pairs of relative value scales based upon different numeraire procedures can be easily transformed--via application of a multiplicative constant--so as to make the

---

*Define the representative charges for the ith and jth medical procedures to be \$P(i) and \$P(j), respectively. As all charges are positive values, if \$P(i) > \$P(j) then \$P(i)/\$N > \$P(j)/\$N. Hence, the selection of numeraire N does not affect the ordinal rankings of the medical procedures on the RVSs. Now, suppose there are two alternative numeraire procedures under consideration with values $\$N_1$ and $\$N_2$ respectively. The value of the Pth procedure on one scale will be $RV_1 = \$P/\$N_1$, while its value on the second will be $RV_2 = \$P/\$N_2$. For any medical procedure, its cardinal value on the second scale is identical to its value on the first, up to a multiplicative constant. That is,

$$RV_1 = k_{12}\ RV_2 \text{ where } k_{12} = \$N_2/\$N_1$$

From the foregoing, we conclude that the choice of numeraire procedure can be made arbitrarily, with no essential effect upon the charge-based RVSs generated.

scales completely comparable.* Consequently, it shall not be necessary to construct separate RVSs for the different surgical, medical, radiology, and laboratory procedure groups of interest in this study.

B. CHARGE-BASED RVSs: PERFORMANCE CRITERIA

Having discussed the construction and conceptual limitations of charge-based relative value scales, it is useful to consider some of the pragmatic issues associated with their development and implementation. In an earlier report in this series, these issues were discussed in detail;** it is useful to review them here for completeness, however.

That earlier work identified eight performance criteria for evaluating RVSs; these were

(1) data required for scale construction;

(2) technical expertise required;

---

*Suppose the set of procedures of interest contains two medical procedures P_m and P_n, and two surgical P_s and P_t. Furthermore, suppose one used medical and surgical-specific numeraire values $\$N_m$ and $\$N_s$ to develop separate Relative Value scales;

$$RV_m = \$P_m/\$N_m < \$P_n/\$N_m$$

$$RV_s = \$P_s/\$N_s < \$P_t/\$N_s$$

The medical-specific and surgical-specific scales could be made comparable (even concatenated) by multiplying one or the other by a constant equal to the ratio of numeraire values. For instance, one could make the surgical scale values comparable to the medical scale by defining

$$k_{sm} = \$N_s/\$N_m$$

and computing

$$k_{sm} \times RV_s = \$P_s/\$N_m < \$P_t/\$N_m$$

This redefined surgical scale is directly comparable to the medical procedure scale (RV_m) defined earlier.

**Hadley, J. et al., "Preliminary Research Plan for Alternative Methods for Developing a Relative Value Scale of Physicians' Fees," The Urban Institute, November 1981, pp. 8-12.

(3) usefulness of scales for pricing problem medical procedures;

(4) potential for future updating of scale values;

(5) feasibility of developing specialty- and location-specific scales;

(6) feasibility of implementation of the scales by individual carriers;

(7) suitability of the scales for the medical procedures under study; and

(8) potential for capturing the external costs of the provision of medical care.

We now assess the performance of charge-based RVSs along these eight dimensions.

Data Requirements

Charge data are potentially available from private insurers, Medicare carriers, and Medicaid programs in every state in the nation. These data generally take the form of claims submitted by physicians. Because of the volume of claims, data processing costs associated with file construction are likely to be high. Charges for each procedure must be arrayed, different percentiles found, and then relativity determined. These steps require multiple passes of large volumes of physicians' claims. Data acquisition costs, however, are relatively low because physicians must submit claims to be reimbursed for their services. Furthermore, most carriers routinely computerize claims in order to process payments. In the long run, total data processing costs may be lower than under the current system, since carriers would not have to compute usual, customary, prevailing, and reasonable charges for each physician and procedure.

<u>Technical Expertise Required</u>

Implementing a charge-based reimbursement system requires substantial computing capacity and sophisticated data processing support. Personnel with advanced training in statistics, economics, or health care financing are not likely to be needed.

<u>Pricing Problem Medical Procedures</u>

There are no particular advantages to a charge-based system in pricing problem procedures. Charges for new procedures may overstate their true marginal costs if the costs of research and development are loaded into the charges for the initial period of performance. They are also unlikely to be a good measure of patients' valuation because of the lack of experience with the procedure as well as, in many cases, insurers' automatic acceptance of the charges set by providers.

Charge-based systems are likely to do a fairly poor job of pricing procedures where the technology is undergoing rapid change. For example, if technological innovations reduce the price of performing certain procedures, it is likely that charges will not fall in response to the lower costs. Because of insurance, prices are likely to remain high, even if costs fall.

<u>Potential for Future Updating of Scales</u>

While data currently exist to develop relative value scales on the basis of physicians' charges, it may be more difficult to implement and update a charge-based RVS if Medicare were to adopt a fee schedule. If the fee schedule were constructed from an RVS based on charges, it could make use of claims (charge) data in the first year. However, subsequent charge data would not be as useful because they most likely would be affected by the allowed fees. Medicare would have to make use of private insurance data as input in updating or altering its RVS in subsequent years, or to rely on methods other than charges for periodic modifications of relative values.

Development of Specialty- and Location-Specific Scales

Charges represent, perhaps, an ideal system for developing different RVSs for different specialties and for different regions or geographic areas. Because charges are available for each physician in each specialty in each region, they represent a mammoth data base from which alternative specialty-specific or locality-specific RVSs can be developed.

Implementation by Individual Carriers

Individual carriers could implement a charge-based method relatively easily. It is unlikely, however, that Medicare would want the carriers to bear this cost because of the large amount of duplication that would occur, especially if relative values are reasonably similar across carrier areas.

Suitability Across Procedures

A charge-based RVS system seems suitable for all classes of procedures. However, it may provide a better guide to relative costs and values within procedure types (that is, within medical, surgical, radiological, and laboratory procedure categories) than across these categories. This is due to the potential biases introduced by the different levels of insurance coverage people have for the various types of procedures. Historically, office visits have not been well insured. Conversely, surgical procedures and diagnostic services, such as laboratory and x-ray services, have been well insured. As a result, office visit fees have not increased at the same rates as fees for surgery, laboratory, and radiology. Thus a charge-based RVS system is likely to overstate the relative prices as well as the relative values of the surgery, laboratory and x-ray procedures and undervalue office visits. This is also true of any other procedures performed that have not been well insured in the past.

External Costs

Charge-based systems do not take account of costs outside the physician's practice, either the costs to patients or costs borne by the hospital. For example, the relative value of a Caesarean section delivery is likely to be greater than that of a vaginal delivery, since physicians' charges are typically greater for the former than the latter. An insurer may wish to "devalue" Caesarean deliveries relative to vaginal deliveries, however, since Caesareans require longer hospital stays and more restricted activity days for the mother than do vaginal deliveries. (The higher relative value of Caesarean sections may also reflect insurance distortions, since it is almost always covered as a hospital procedure, while vaginal deliveries are generally less well covered.) These concerns are relevant if HCFA is concerned about total costs rather than only payments to physicians.

Summary and Implications for Assessing Charge-Based RV Scales

As discussed, charge-based RVSs are attractive for a number of reasons. Chief among these are the general availability of requisite data and the ease with which the scales can be constructed. Also they are easily understood and employed by individuals possessing substantive knowledge of medical care markets but who may lack advanced training in economics or statistics. Furthermore, charge-based RVSs can be implemented rather easily by carriers, and they are readily adapted to reflect regional or specialty distinctions across procedures.

The primary weakness of charge-based RVSs is the potential bias imparted to them by imperfections in the market place including imperfect information and the distorting effect of health insurance upon individuals' perceptions of the true price of medical care. Furthermore, these distortions and imperfections may vary by carrier, by geography, or by specialty of the physician providing the service under investigation. Therefore, RVSs may be sensitive to the data bases underlying them.

As part of this evaluation, different versions of charge-based RVSs are constructed over alternate data bases employing different measures of representative charges. Similarities across the scales (in terms of the ordinal and cardinal rankings of particular medical procedures) can be interpreted as evidence that any bias in the RVS values is a consistent and general phenomenon and not specific to particular market areas or classes of medical procedures.

Furthermore, the possible biases inherent in charge-based RVSs suggest the need for either subjective, expert assessment or some type of independent cost-finding. The physician-consultant associated with this project will examine the values assigned to a set of medical procedures by the several RVSs and will identify any which appear to deviate from his own subjective scale of value.

C. METHODOLOGY FOR AN EMPIRICAL COMPARISON OF CHARGE-BASED RVSs

In this section the overall approach to the development of charge-based relative value scales is discussed. First is a description of the four data bases which are, because of their accessibility and content, the primary sources of information employed to construct the charge-based scales. Following that are the methods for constructing measures of representative charges for these procedures. Finally, the sets of medical procedures evaluated along the different RVSs are identified.

Available Data

Four sources of data on physicians' charges are readily accessible and can be employed to develop RVSs for various sets of medical procedures. The first is the annual data on local prevailing and customary charges provided by Medicare carriers, nationwide, to the Health Care Financing Administration (HCFA). The next two are samples of Medicare and Medicaid claims collected by The Urban Institute from California physicians during the first quarter of calendar years 1978 and 1974. The fourth file is the Health Insurance Association of America's (HIAA's) survey of surgical claims obtained from over twenty commercial insurers across the country.

The primary sources of data for each file are physicians' claims for reimbursement for medical services performed over some relevant time period. The files differ in terms of the types of medical services represented therein, the geographic and temporal distribution of the basic claims data, and the types of charge information available. We shall briefly discuss these characteristics of each of the three data bases in sequence.

HCFA Prevailing Charge Survey

Each year, Medicare intermediaries representing over 200 Medicare reimbursement localities (states or portions of states nationwide) provide HCFA with local data on physicians' charges for a set of procedures provided, with high frequency, to Medicare beneficiaries. The most recent available data are from Fee Screen Year (FSY) 1982,* and represent actual charges billed during calendar year 1980. Medicare intermediaries provided data on the local

---

*Fee Screen Year 1982 spans the period July 1981 through June 1982. Allowed Medicare reimbursements for services performed during FSY 1982 are dependent, in part, upon regional average charges for the service performed (and billed) during the preceding calendar year, in this case 1980.

prevailing charge for 103 high-volume procedures as well as the median (50th percentile point) and unindexed 75th percentile point of the distributions of local customary fees for these procedures.* Prevailing fees and percentile points on regional distributions of customary fees are specific to the particular physician-specialties most likely to have performed the procedure.

The Urban Institute California Medicare-Medicaid Claim File

These files are based upon Medicare (and joint Medicare-Medicaid) claims submitted by a sample of over 7,000 California physicians during the first quarter of each of the five years 1974-1978. Of special interest are charge data for particular medical procedures performed by each physician during the earliest and latest years, 1974 and 1978. We shall construct RVSs based upon these charge data; for 1978 the file contains information on approximately 139,000 physician-procedure combinations representing eleven medical specialties. Over half the physician-procedure combinations are for physicians in five specialties: General Practice, Internal Medicine, General Surgery, Orthopedic Surgery, and Ophthalmology.

For each physician-procedure combination the file contains data on the physician's average charge, the average amount reimbursed by Medicare, and the

---

*Prior to 1976, the regional prevailing charge for a procedure was defined as the 75th percentile point on the regional distribution of physicians' customary fees for that procedure. (Customary fees are defined as a physician's median charge in the preceding year for the particular procedure). Beginning in 1976, indexed prevailing fees were constrained to be the lower of the current year's regional prevailing fee or an adjusted prevailing fee defined to be the product of a base-year (1973) prevailing fee and the current year's value of the Medicare Economic Index.

frequency with which the physician performed the procedure.* The files also contain data on the local prevailing charge and the physician's customary charge for each procedure.

HIAA Surgical Prevailing Charge File

This file is derived from annual surveys of 22 large commercial insurance carriers conducted by the Health Insurance Association of America (HIAA); the data we possess cover the time period September 1977 through August 1978.

From these surveys, HIAA computes several measures of the average surgical procedure charges in 250 geographic areas around the country. For each of approximately 160,000 procedure-region pairs, HIAA computes mean and modal charges; for regions where at least five claims were filed during the year, HIAA also reports the values of different percentile points along the regional distributions of charges for each procedure.**

Data Summary

As RVSs may be sensitive to the sources of the data upon which they are based, we identified several sources of charge data for use in developing alternative scales. By contrasting the cardinal measures (and rank order) assigned to medical procedures by each relative value scale, it is possible to evaluate charge-based scales for consistency. If consistency across scales is established, then a case can be made for selecting the data on the basis of cost and availability rather than on conceptual grounds.

---

*Average charges are the service-weighted averages of the physician's Medicare and Medicaid billings for that procedure. Average reimbursements are the average amount allowed by Medicare for that service performed by the physician.

**For the area-procedure pairs for which at least 5 claims were processed, HIAA reports the 50th, 60th, 70th, 75th, 80th, 90th and 95th percentile charges.

Each of the identified data files differs in terms of the charge data contained therein, the services represented, and the time frame and geographic region covered. The HCFA prevailing charge file is relatively up-to-date and nationwide in scope. It is, however, limited in terms of the medical procedures included. The Urban Institute's California Medicare Claims files are much broader in terms of the set of procedures covered. Furthermore, as the Urban Institute files are based upon actual claims submitted for reimbursement purposes, the data are of especially high quality. This information is somewhat dated, however, and limited to a single, albeit very large, state. The HIAA data are limited to surgical procedures but (in conjunction with all other data) should prove useful for contrasting relative prices for procedures reimbursed through public insurance programs with charges paid by private carriers.

Measures of Representative Charges for Medical Procedures

In a previous section, the RVS value for the pth medical procedure was defined to be RVS_p = \$P/\$N where \$P and \$N are representative charges for the pth procedure and the numeraire, respectively. In this section we discuss the methods for developing these representative values from distributions of charge data available from the sources we identified.

Recall that our data are at the physician or region unit of observation; associated with each physician or region are charge data for a set of medical procedures (to facilitate discussion, these alternative charge measures are arrayed in Table 2.1).

Table 2.1

Charge Data by Data Base and Unit of Observation

	File Name	Unit of Observation	Types of Available Charge Data
(1)	HCFA Prevailing Charge File	Region-procedure combinations	(a) adjusted prevailing charge (b) (unadjusted 75th percentile point of distribution of local customary fees (c) 50th percentile point of distribtion of local customary fees
(2)	UI California Medicare Claims File: 1974 and 1978	Physician-procedure combinations	(a) average Medicare and Medicaid charges (b) arithmetic mean of Medicare allowed reimbursements (c) Physician's customary fee
(3)	HIAA Surgical Prevailing Charge File	Region-procedure combination	(a) regional mean charge (b) regional modal charge (c) various percentile points on regional distribution of charges

For each charge type associated with a particular medical procedure, there exists a distribution of values across the regions (or physicians) on the data file. Rather than identifying a single value to be uniquely "representative" of the charge distributions, an array of alternative representative values is developed. For example, as shown in Table 1, the regional prevailing fee is one of the charges associated with each procedure on the HCFA data file. From the distribution of these prevailings over all the regions on the file, four points are selected as alternative "representative" values. These are: 1) the mean, 2) the median (50th percentile point), 3) the 75th percentile point, and 4) the 90th percentile point on the distribution. The mean and the median values are common measures of the central tendency of any distribution. Two points above the median charge were also selected, since local Medicare prevailing charges are based upon values at the upper end of regional charge distributions.

From the four representative values on the distributions of each of the three types of charges on the HCFA file, 12 alternative relative value scales can be constructed.* The same method is followed for identifying alternative representative charge values for the procedures on The Urban Institute's California claims files and the HIAA Surgical Procedure Charge file. For each type of charge data on these files, the same four points on the distributions of charges are selected to be alternative representative values; and RVSs are computed from representative charges of procedures and the numeraire.

---

*A 13th representative charge and scale can be developed from the 1969 California Relative Studies (CRVS) value associated with each procedure. Given a procedure-group-specific conversion factor CF_j--which is a measure of the dollar value per CRVS unit of procedures in the jth group--the alternative scale (CRVU) is constructed as follows. $CRVU_i = (CRVS_i \times CF_j) \div (CRVS_n \times CF_n)$ where i denotes the ith procedure and n denotes the numeraire procedure.

D. COMPARISON OF SCALES: ALTERNATIVELY DEFINED REPRESENTATIVE CHARGE VALUES

If charge-based RVSs are to be useful to policy makers, third party payers, and others, they should be robust with respect to the source of the data upon which they are based. If RV scales are to be the foundation of physician reimbursement schedules, they must be stable across time and regions. In this section, we address one dimension of this issue: the sensitivity of RVSs to the definition of the procedure-specific "representative" charge. Using a national cross-section of Medicare charge data--the HCFA Prevailing Charge file described previously--we compare and contrast a set of relative value scales constructed from alternative representative charge values. We also perform a similar analysis of scales developed from The Urban Institute's 1978 California data base. The HCFA file, besides being national in scope, is representative of the types of medical procedures frequently provided to Medicare beneficiaries. Hence, the assessment of the stability of these procedures' relative values on alternatively constructed scales should be especially interesting to HCFA staff. The Urban Institute file is much more comprehensive in terms of the procedures represented and affords us the opportunity to study alternatively constructed RVSs in more detail, as we shall subsequently demonstrate.

Comparing RVSs Developed from HCFA Prevailing Charge File Data

For each of the 103 procedures on the HCFA file listed in Table 2.2, we computed twelve relative values; each value is associated with a different RV scale.* In Table 2.3 are simple mnemonic names for each scale, along with the type-of-charge and point on the distribution of charges which are the basis of the RVSs.

While any of the 103 procedures could serve as numeraire, we selected the procedure performed most frequently according to the HCFA data--a Brief Hospital Visit for an Established Patient. Consequently, scale values are in units of hospital-visit-equivalents, the numeraire having a value of 1.0 on the scale.

Following assignment of relative values to each procedure on the 12 RVSs, we examine the scales for consistency. Simple descriptive statistics corresponding to each scale are presented in Table 2.4; following that, in Table 2.5 are Pearson product-moment correlation coefficients (r) for pairs of RVSs; nad Spearman rank-order correlations (r_s) for pairs of scales are

---

*As we wished to construct our RVSs from data on local distributions of charges, we did not want to include among these data any observations from regions where physician reimbursements were based upon fee schedules. Therefore, we eliminated from our analysis records where the source of the locally prevailing charge was coded as F (relative value derived), G (other method), or J (an unspecified code which erroneously appeared on some records in the HCFA file). We also eliminated records from the state of North Dakota, as prevailing charge data on these records proved to be unreadable due to problems with their format.

Furthermore, we defined as missing those records for which both the 50th and 75th percentile values of the customary charge distribution had values less than 0. A small number of records also had unreadable data in the 75th percentile charge field; we did not use these data in computing RVSs based upon 75th percentile points of unadjusted charge distrubtions.

Finally, on approximately 15 percent of the records the charge type code identified the prevailing charge as representing the "professional component" only; unless otherwise indicated in the text, we did not employ these data when computing charge-based RVSs.

Table 2.2

Terminology Used in the Medicare Directory of Prevailing Charges

Due to studies of high volume medical procedures conducted in the past, only 30 of the 103 procedures will have prevailing charge screens recorded for both General Practitioners and selected Specialists. The remaining procedures will have charge screens relating only to the category of specialist specified below:

Procedure	Terminology	Specialist(s)
1	Initial Brief Office Visit	Internist
2	Initial Limited Office Visit	GP, Internist
3	Initial Intermediate Office Visit	Internist
4	Initial Comprehensive Office Visit	GP, Internist
5	Minimal Follow-up Office Visit	GP, Internist
6	Brief Follow-up Office Visit	GP, Internist
7	Limited Follow-up Office Visit	GP, Internist
8	Intermediate Follow-up Office Visit	GP, Internist
9	Extended Follow-up Office Visit	GP, Internist
10	Comprehensive Follow-up Office Visit	GP, Internist
11	Brief Follow-up Office Visit	GP, Internist
12	Limited Follow-up Home Visit	Family Practitioner
13	Intermediate Follow-up Home Visit	GP, Family Practitioner
14	Extended Care Facility Visit	Internist
15[b]	Brief Follow-up Nursing Home Visit	GP, Internist
16	Initial Brief Hospital Visit	GP, Internist
17	Initial Intermediate Hospital Visit	Internist
18	Initial Comprehensive Hospital Visit	GP, Internist
19	Brief Follow-up Hospital Visit	GP, Internist
20	Limited Follow-up Hospital Visit	GP, Internist
21	Intermediate Follow-up Hospital Visit	GP, Internist
22	Extended Follow-up Hospital Visit	Internist
23	Brief Emergency Room Visit	Internist
24	Limited Emergency Room Visit	Internist
25	Intermediate Emergency Room Visit	Internist
26	Limited Consultation	GP, Internist
27	Extensive Consultation	Internist
28	Comprehensive Consultation	Internist
29	Psychotherapy-One Hour	Psychiatrist
30	Psychotherapy-Half Hour	Psychiatrist
31[b]	Chiropractic Office Visit	GP, Chiropractor
32	Initital Physiotherapy	Physical Therapist
33[b]	Follow-up Podiatric Office Visit	Podiatrist
34	Electrocardiogram (EKG)	GP, Internist
35	EKG-Interpretation, Report Only	GP, Internist
36	Spirometry	Pulmonary Diseases
37	Electroencephalogram (EEG)	Neurologist
38	Chemotherapy	Internist
39[b]	Collection of Specimens	Laboratory[a]
40[b]	Debridement of Nails	Surgeon
41	Skin Biopsy	Surgeon
42[b]	Chemotherapy	Internist
43	Radical Mastectomy (Modified)	Surgeon
44	Open Reduction of Fracture	Orthopedic Surgeon
45	Arthrocentesis-Major Joint	GP, Orthopedic Surgeon
46[b]	Coronary By-Pass (Three or More Arteries)	Orthopedic Surgeon
47	Artificial Hip Replacement	Orthopedic Surgeon
48	Needle Puncture of Bursa	Surgeon
49	Bronchoscopy	Cardiologist
50	Thoracentesis	Cardiologist
51[b]	Catherization of Heart	Cardiologist
52[b]	Insertion of Pacemaker	Surgeon
53	Partial Colectomy	General Surgeon
54	Appendectomy	GP, General Surgeon
55	Sigmoidoscopy	General Surgeon
56[b]	Hemorrhoidectomy	General Surgeon
57	Cholecystectomy	General Surgeon

Table 2.2 (Continued)

Terminology Used in the Medicare Directory of Prevailing Charges

Procedure	Terminology	Specialist(s)
58	Repair Hernia	General Surgeon
59[b]	Diagnostic Cystourethroscopy	GP, Urologist
60[b]	Dilation of Urethra	Urologist
61[b]	Prostatectomy	Urologist
62[b]	Electrosection-Prostate (TUR)	GP, Urologist
63	Hysterectomy	Obstetrician-Gynecologist
64	Initial Complete Eye Exam	Opthalmologist
65	Comprehensive Eye Exam	Opthalmologist
66	Eye Exam with Tonometry	Opthalmologist
67	Extraction of Lens	GP, Opthalomologist
68	Chest X-ray-Single View	GP, Radiologist
69	Chest X-ray-Two Views	GP, Radiologist
70	X-ray-Spine	Radiologist
71	X-ray-Hip	Radiologist
72	X-ray-Upper GI Tract	GP, Radiologist
73	X-ray-Colon	Radiologist
74	Radiation Therapy-Low Volt	GP, Radiologist
75	Radiation Therapy-Super Volt	GP, Radiologist
76	Radiation Therapy-Megavolt	Radiologist
77[b]	CAT Scan-Head	Radiologist
78[b]	CAT Scan-Abdomen	Radiologist
79	Three Chemistry Tests	Laboratory[a]
80[b]	Nineteen Chemistry Tests	Laboratory
81	Culture-Other than Blood	Laboratory
82	Hemoglobin	Laboratory
83	Automated Blood Count	Laboratory
84	White Cell Count	Laboratory
85	Complete Blood Count (CBC)	Laboratory
86	Cholesterol Test	Laboratory
87	Flocculation Test	Laboratroy
88	Hematocrit	Laboratory
89	Platelet Count (Rees-Ecker)	Laboratory
90	Potassium Test	Laboratory
91	Prothrombin Time Test	Laboratory
92	Sedimentation Rate	Laboratory
93	Blood Sugar	Laboratory
94	BUN-Urea Nitrogen	Laboratory
95	Uric Acid	Laboratory
96	Feces-Occult Blood	Laboratory
97	Pap Test	Laboratory
98	Routine Urinalysis	Laboratory
99	Chemical Urinalysis	Laboratory
100	Pathology-Three Specimens	Pathology
101[b]	Monitoring of Pacemaker	Cardiologist
102[b]	Donor Nephrectomy	Urology
103[b]	Kidney Transplant	Urology

Source: Medicare Directory of Prevailing Charges 1981, U.S. Department of Health and Human Services, Health Care Financing Administration, and unpublished HCFA data.

Notes: a. Without regard to medical specialty or place where services was performed.

b. Procedure-specialty combinations for which data were not available on The Urban Institute's California Medicare Charge File.

Table 2.3

Descriptions of Relative Value Scales
Constructed from HCFA Prevailing Charge File

RV Scale Name	Charge Type	Point on Charge Distribution Selected as Representative
MNP	adjusted prevailing	Mean
MN50	50th percentile of customaries	
MN75	75th percentile of customaries	
MDP	adjusted prevailing	Median
MD50	50th percentile of customaries	(50th percentile)
MD75	75th percentile of customaries	
UQP	adjusted prevailing	75th percentile pt.
UQ50	50th percentile of customaries	(upper quartile)
UQ75	75th percentile of customaries	
UDP	adjusted prevailing	90th percentile pt.
UD50	50th percentile of customaries	(upper decile)
UD75	75th percentile of customaries	

Table 2.4

RV Scales--Descriptive Statistics
(HCFA Prevailing Charge File)

Scale Name	n	Mean	Median	St. Dev.	C.V.
MNP	103	13.30	1.80	31.47	2.37
MN50	103	12.67	1.63	29.10	2.30
MN75	103	12.29	1.62	28.44	2.31
MDP	103	13.49	1.79	32.60	2.42
MD50	103	12.83	1.56	30.42	2.37
MD75	103	11.86	1.50	28.56	2.40
UQP	103	13.03	1.84	31.26	2.40
UQ50	103	11.72	1.50	27.60	2.36
UQ75	103	12.58	1.64	29.74	2.36
UDP	103	14.32	2.14	32.91	2.30
UD50	103	13.98	1.78	32.09	2.30
UD75	103	12.40	1.64	28.33	2.28

Table 2.5

Pearson Correlation Coefficients*

(103 procedures on RVSs developed from HCFA Prevailing Charge Data File)

	MNP	MN50	MN75	MDP	MD50	MD75	UQP	UQ50	UQ75	UDP	UD50	UD75
MNP	1.000	.998	.998	.999	.999	.997	.999	.998	.997	.997	.994	.996
MN50		1.000	.999	.994	.997	.994	.995	.997	.995	.997	.997	.999
MN75			1.000	.995	.995	.997	.995	.997	.998	.996	.993	.997
MDP				1.000	.995	.997	.999	.997	.995	.994	.989	.992
MD50					1.000	.995	.998	.997	.993	.995	.994	.995
MD75						1.000	.995	.997	.997	.991	.985	.991
UQP							1.000	.998	.995	.996	.993	.994
UQ50								1.000	.998	.998	.994	.997
UQ75									1.000	.996	.990	.995
UDP										1.000	.997	.998
UD50											1.000	.998
UD75												1.000

*All correlations are significantly greater than zero at the .0001 level.

Key:
MNP - mean adjusted prevailing charge
MN50 - mean of median customaries
MN75 - mean unadjusted prevailing
MDP - median of adjusted prevailings
MD50 - median of median customaries
MD75 - median of unadjusted prevailings
UQP - upper quartile of median customaries
UQ50 - upper quartile of median customaries
UQ75 - upper quartile of unadjusted prevailings
UDP - upper decile of adjusted prevailings
UD50 - upper decile of median customaries
UP75 - upper decile of unadjusted prevailings

reported in Table 2.6. The Pearson correlations measure the degree of linear relationship between two scales, while the Spearman coefficients measure the similarity in the rankings of procedures on different RVSs.

One method for assessing the overall similarity of scores on the different scales is to compute the mean of the scores of the 103 procedures on each of the RVSs. To the extent that procedures' scores on pairs of scales differ in a uniform manner--all scores on one scale being greater or smaller than corresponding scores on the other scale--the means will provide a rough measure of these differences. As reported in Table 2.4, mean RVS values for a constant basket of medical services vary between 11.72 and 14.32, a difference of 22 percent. Because the scores of procedures on the HCFA file were not symmetrically distributed about the mean score on the 12 RVSs, we also report the median value on each scale.* These ranged between 1.50 and 2.14, a difference of 43 percent. This variability in both measures of average score for a fixed basket of services suggests that procedures' scores do differ somewhat across the scales.

The significance of this finding vis-a-vis the usefulness of charge-based scales for the development of fee schedules is not entirely clear. If, for example, the ratios of scores across pairs of RVSs are approximately constant for all procedures, then the variation in score levels across scales would be irrelevant. If the RVSs are to be the bases of fee schedules, differences in scale values could be eliminated via the application of an appropriately constructed monetary conversion factor. For example, two scales, one with a mean of 10 and the other a mean of 5--for a common set of procedures--would

---

*On each of the 12 scales, over half of the 103 procedures had relative values of 2.00 or smaller. At the same time, approximately 10 percent had RVS values in excess of 50.0. This set of HCFA procedures was skewed toward the low end.

Table 2.6

Spearman Rank-Order Correlations*
(103 procedures on RVSs developed from
HCFA Prevailing Charge Data File)

	MNP	MN50	MN75	MDP	MD50	MD75	UQP	UQ50	UQ75	UDP	UD50	UD75
MNP	1.000	.996	.995	.996	.994	.994	.995	.988	.991	.994	.992	.993
MN50		1.000	.997	.990	.995	.990	.986	.990	.991	.988	.991	.990
MN75			1.000	.990	.991	.992	.988	.987	.993	.990	.991	.994
MDP				1.000	.993	.995	.996	.986	.990	.992	.990	.991
MD50					1.000	.994	.991	.990	.991	.988	.992	.988
MD75						1.000	.994	.989	.994	.991	.991	.992
UQP							1.000	.990	.992	.995	.991	.991
UQ50								1.000	.994	.987	.992	.989
UQ75									1.000	.993	.995	.996
UDP										1.000	.994	.995
UD50											1.000	.996
UD75												1.000

*All correlations are significantly different from zero at the .0001 level.

Key: MNP - mean adjusted prevailing charge
MN50 - mean of median customaries
MN75 - mean unadjusted prevailing
MDP - median of adjusted prevailings
MD50 - median of median customaries
MD75 - median of unadjusted prevailings
UQP - upper quartile of median customaries
UQ50 - upper quartile of median customaries
UQ75 - upper quartile of unadjusted prevailings
UDP - upper decile of adjusted prevailings
UD50 - upper decile of median customaries
UP75 - upper decile of unadjusted prevailings

generate identical fee schedules if the conversion factor for the fist were $5 per unit and the conversion factor for the second $10 per unit.

Besides computing measures of the average value of the HCFA procedures on the alternative RVSs, we also calculated coefficients of variation (C.V.), which measure the spread of the 103 procedures' scores about the mean score on each scale.* As reported in Table 2.4, coefficients of variation are relatively constant across the scales. Only a six percent difference separates the largest and smallest C.V.s, 2.42 vs. 2.28. Standard deviations, likewise, are similar, ranging from 27.10 to 32.91, a difference of 19 percent between the largest and smallest. It is evident, therefore, that the relative values of the 103 procedures represented on the HCFA file are similarly distributed about the mean value on all 12 scales.

In sum, the scores of particular procedures vary from scale to scale, but the dispersion of scores--for a fixed basket of services--is relatively constant across the scales. This suggests that values on one RVS may simply be multiples of their values on another; that is, it suggests a possible linear relationship between RVSs. To assess the extent to which RVSs are

---

*The coefficient of variation (C.V.) is the ratio of the standard deviation to the mean of a particular distribution. Effectively, the C.V. is a measure of the spread of a distribution adjusted for the value of its mean and is a useful statistic for contrasting distributions whose means differ.

If two RVSs are linearly related, then scores of particular medical procedures on the scales will differ by a multiplicative constant. For instance, if $RV_2 = cRV_1$--that is, if the two scales are linearly related--the mean value of a set of procedures on the first scale is $\overline{RV}_1$, while the mean value on the second will be $\overline{RV}_2 = c\overline{RV}_1$. Standard deviations (s) will also differ by this constant, being s_1 and $s_2 = cs_1$ on the first and second scale, respectively. On the other hand, coefficients of variation will be identical, since $s_1/\overline{RV}_1 = cs_1/c\overline{RV}_1 = s_2/\overline{RV}_2$. Therefore, the C.V. may be the more appropriate measure of dispersion when the objective is to compare RVSs which are linearly related.

linearly related, it is necessary to examine the Pearson correlation coefficients.*

The data of Table 2.5 confirm the existence of strong and statistically significant linear relationships among the alternative RVSs. Correlations between scores of the HCFA procedures on different scales range from .96 to .99 in value. Such large r values--relative to the maximum value 1.0--indicate that procedures' scores on any scale are essentially equal (up to a multiplicative constant or scale factor) to scores on the others. Furthermore, in all case, these high correlation coefficient values are significantly different from zero at the .0001 level.**

An alternative interpretation of these data can be developed by computing squared values of the correlation coefficients. This new statistic is precisely the familiar r-square term traditionally reported in least squares regression analyses; it is a measure of the variation (about the mean) observed on one set of RVS scores, that is "explained" by variation in scores

---

*If two scales are linearly related, then the least squares regression equation summarizing this relationship is $RV_1 = b_{12}\ RV_2$. By construction, the correlation coefficient (r) is a function of the slope of the least squares line and the standard deviations (s) of scales of medical procedures evaluated on each scale: $r = b_{12}(s_2/s_1)$.

As Blalock (Social Statistics, 1972, pp. 376-378) explains, the correlation coefficient can be interpreted as a "measure of the goodness of fit of the least-squares straight line." High values of r--close to 1.0 in absolute value--are indicative of tight distributions of observations about the least squares regression line.

**The significance level is interpretable as the probability that one would observe the sample correlation coefficient (given in the table) under the condition that the true population correlation is actually zero; these probabilities are based on an F statistic computed as

$$F = \frac{r^2}{(1-r^2)}(N-2)$$

with (1) and (N-2) degrees of freedom.

on a different scale about their mean.* The observed values of correlation coefficients--uniformly .99 or greater--correspond to squared values of .99 as well. In other words, well over 90 percent of the observed variation in the scores of medical procedures measured along any one of the 12 scales is "explained" by the corresponding variation in scores on any one of the other scales.

Similarities in terms of the rankings of procedures on the 12 RVSs can be determined from the Spearman rank-order correlation coefficients (r_s) found in Table 2.6. Essentially r_s can be interpreted as the product-moment correlation between the rankings of medical procedures on two relative value

---

*Following Blalock (1972, p. 392), define the least-squares relationship between two variables (e.g., between two relative value scales) to be

$$\hat{RV}_2 = b\ RV_1$$

Furthermore, define the means of the two sets of variables to be

$$\overline{RV}_2 \text{ and } \overline{RV}_1, \text{ respectively.}$$

Blaloch then demonstrates that

$$r^2 = \frac{(\hat{RV}_2 - \overline{RV}_2)^2}{(RV_2 - \overline{RV}_2)^2} = \frac{b^2\ (RV_1 - \overline{RV}_1)^2}{(RV_2 - \overline{RV}_2)^2}$$

Hence, r^2 is interpretable as the fraction of the variation in observed values on any one scale (RV_2) about its mean, explained by the observed variation in corresponding measures about the mean of another scale (RV_1). As Blalock (p. 391) relates, "By _explained_, of course, we do not imply a causal explanation but merely an association between the two variables."

scales.* In Table 2.6, values of r_s are shown to be consistently greater than .98; many exceed .99 in magnitude. As might have been expected, given the high degree of correlation (r) between procedures' cardinal scores on the different RVSs, there is considerable agreement across the scales in terms of procedures' ordinal rankings as well.

Comparing RVSs Developed from The Urban Institute's 1978 California Data File

We created RVSs from The Urban Institute's file of Medicare and Medicaid claims generated by California physicians during the first quarter of 1978. On the file are physicians' average billings, average reimbursements, and customary charges for 443 medical procedures. We developed relative value scales for each procedure in a manner analogous to the approach we employed for the HCFA scales. We created individual scales by first choosing a particular charge type (e.g., average billings) then defining the "representative" charge for each procedure to be one of four points--mean, median, 75th percentile, 90th percentile--on the distribution of charges across all sampled California physicians. Relative scale values are the ratio of procedures' representative charges to the representative charge of a numeraire--in this

---

*Blalock (1972, p. 416) described the construction of the Spearman rank order correlation coefficient r_s. Paraphrasing Blalock, let D_i be the difference of ranks on two scales for any observation i; squaring these D_i values, summing the results (and constraining the sum to be within the range ± 1.0) we obtain:

$$r_s = 1 - \frac{6 \sum_{i=1}^{N} D_i^2}{N(N^2-1)}$$

An r_s value of 0.0 implies the absence of any rank-order correlation; extreme (+1 or -1) values imply perfect agreement or disagreement of rankings along two scales.

case a simple surgical procedure, Needle Puncture of Bursa.* The three charge types and four points on the charge distributions enabled us to construct 12 representative charges (and 12 relative value scales) per procedure.**

The analysis of these 12 RVSs followed the same pattern established during our assessment of scales based upon HCFA data. We first calculated Pearson product-moment correlations of the procedures' scores on each pair of scales. In all cases, these correlations equaled or exceeded .99 in value and were statistically significant at the .0001 level. We also computed Spearman rank-order correlations of the procedures' scores. Here, too, values exceeded .95 and were significant statistically at the .0001 level.

Given the large number of procedures on the 1978 California file, we were able to contrast the RVS scores of procedures in particular type-of-service groups. On the file were 89 radiological, 142 surgical, 93 pathology, and 119 medical procedures. Looking at procedures in each group separately, we computed Spearman and Pearson correlations of their scores on the different scales. In all cases, these correlations were statistically significant at the .0001 level. The magnitude of the coefficients did vary somewhat by procedures' service type. For medical and radiological procedures, scores on relative value scales had Pearson correlations in the range .92 to .99, while

---

*As we discussed in an earlier section, any procedure could serve as numeraire. We did not use the same procedure--Hospital Visit for an Established Patient--that we employed as numeraire when constructing scales from the HCFA data because it does not exist on all the data files available as part of this study.

**In 1978, California Medicare carriers did not compute physicians' customary fees for pathology procedures in the usual manner, i.e., from charge data. Rather, a <u>de facto</u> fee schedule approach was employed whereby customary fees were determined from the CRVS value for pathology procedures. For this reason, we did not compute charge-based RVSs for pathology procedures from customary fee data.

surgical procedures' scores were correlated at the .99 level on all pairs of RVSs. Pathology procedures' scores exhibited the greatest range in r values--between .86 and .99. The Spearman correlations between procedures' scores on pairs of scales were somewhat lower for medical and pathology procedures than for radiology. In the former cases, r_s values ranged between .89 and .99, while for radiology and surgery, procedures' RVS scores were (Spearman) correlated in the range .97 to .99 across the scales.

Given these results, we conclude that relative values of medical procedures on charge-based scales are robust with respect to variation in the type of charge data underlying the scale. Procedures' scores were highly correlated on RVSs developed from alternatively defined "representative" charges; this was true of scales developed both from HCFA prevailing charge data as well as from The Urban Institute's sample of 1978 California Medicare claims. Pathology and, to a lesser degree, medical procedures' scores on alternative RVSs exhibited somewhat greater variation--as measured by smaller correlation coefficient values--than did radiology and surgical procedures. Even so, these small correlations were still substantial in absolute terms, being above .85 in value in all cases. Therefore, our conclusion of robustness of RVS values across types of charge data holds for all procedures as well as for procedures grouped into specific service categories.

E. COMPARISON OF SCALES BASED UPON ALTERNATIVE SOURCES OF CHARGE DATA

In the foregoing analyses, we demonstrated that RVS values were robust with respect to the type of charge data--billings, reimbursements, customary fees, etc.--underlying them. It remains to be determined whether or not these values are similarly robust across scales based upon data from different geographic regions or at different times.

In this section, we contrast five RVSs constructed from alternative data bases. Two of these scales have previously been identified. The first is based upon the HCFA Prevailing Charge file. The representative charge for each procedure is the 75th percentile of the distribution of median local customary fees.* The second scale is based upon billings data obtained from The Urban Institute's 1978 California sample of physicians' Medicare claims. On this scale, representative charges are 75th percentile points on the distribution of physicians' average billings for each procedure. We develop a third scale in the same fashion using billings data from the 1974 Urban Institute Physicians' Medicare Claims file. These latter two scales are 78CAL and 74CAL, respectively.

We constructed the fourth RVS from the nationwide survey of surgical prevailing charges conducted by the Health Insurance Association of America (HIAA). Among other data, that file contains 1978 median billings for surgical procedures charged by physicians in nearly 250 regions of the country. We created a relative value scale using 75th percentiles on the distributions of these medians as "representative" values.

The fifth scale is not precisely a charge-based RVS in the mold of the other four. Rather, it is a provider group's scale of procedure values possessing only an indirect link to historical market prices for physicians' services. This scale, the Mountain Medical Affiliates Relative Value Study (MMA-RVS) is based upon the 1971 Colorado Medical Society's relative value

---

*This is the RVS identified in an earlier section as UQ50.

scale--which MMA-RVS was intended by its designers to replace.* Associated with procedures on this scale are unit values representing their relative worth vis-a-vis other procedures in the same service category. By multiplying these unit values by category-specific Conversion Factors (supplied by MMA) we obtained a set of relative values on a common scale.** While MMA-RVS is not a true charge-based scale, it affords us the opportunity to contrast our four other RVSs with a readily available alternative.

In this chapter, we are especially interested in the HCFA Prevailing Charge file as a source of data, since it contains information on high-volume Medicare procedures. If we demonstrate the robustness of charge-based RVSs with respect to the values of these procedures, we will have helped establish the validity of these scales as potential bases for the development of Medicare fee schedules.

As high-volume Medicare procedures are of special interest to us, we constructed relative values for each of the 103 HCFA procedures represented on the five data sources discussed above. While charge data on all 103

---

*In the Preface to the MMA Relative Value Study, its authors state that "The unit values that were developed by the Colorado Relative Value Committee in 1975 and 1976 were used as reference. In addition, many unit values in this study have been revised and added due to recent procedural and technological advances in the field of medicine." As these earlier Colorado scales were based, in part, upon the still-earlier California Relative Value Studies, the Colorado scales have some relationship to California physicians' charges. However, given the modifications and adjustments to the California scale performed in each stage of the Colorado scale's development, the precise relationship between historical charges and the current version of the MMA-RVS is not clear.

**On the MMA-RVS, relative scale values are provided for medical procedures in each of five categories. MMA also provide a set of conversion factors which transform procedure group-specific scale values to a common dollar metric. In other words, the product of MMA SCALE VALUE x CONVERSION FACTOR for each procedure yields a dollar value for that procedure. Because MMA developed and made available to us these conversion factors, we use the MMA scale in this comparative analysis rather than the more widely known California Relative Studies values.

procedures were not available on each file, a sufficiently large number was present in each case to permit cross-scale analyses of relative values.* Consistent with the approach we have employed throughout this chapter, procedures' relative values are the ratios of their representative charges to that of the numeraire--Needle Puncture of Bursa.

In Tables 2.7a and 2.7b are Pearson r and Spearman r_s correlations of procedures' relative value scores on the five RVSs. In all cases, these product-moment and rank-order correlations were quite large--exceeding .94 in value. From these data we conclude that the relative values (and relative rankings) of frequently performed Medicare procedures are similar--i.e., are strongly linearly related--across the five RVSs. As each of the charge-based scales yields analytically equivalent values and rankings for a set of high-volume Medicare procedures, any of the scales might reasonably be used as the basis of a fee schedule. However, as some data bases (notably the HCFA Prevailing Charge file and the two Urban Institute files) are more representative of Medicare procedures than are the others, these may be the more useful in the context of developing new Medicare reimbursement schemes.

---

*Of the 103 procedures of the HCFA Prevailing Charge file, 21 were present on the HIAA Surgical file, 95 on the MMA-RVS and 83 on The Urban Institute's 1978 California file. Of the latter, we also obtained data on 82 from The Institute's 1974 California file. While the two Institute files and the HCFA file are supposedly representative of high-volume Medicare procedures, not all procedures on the latter were also available on the former two. The reason for this is two-fold. First, Medicare claims from physicians in only 11 specialties are represented on The Urban Institute's files; Cardiology, Podiatry, Neurology, Physical Therapy, and Urology were not among these 11 (but are among the specialties whose charge data constitute prevailing fees for one or more procedures on the HCFA file). Second, procedures such as Head and Abdominal CAT Scans may be high frequency procedures in the 1980s but may not have been so common in the mid-1970s, the period for which The Urban Institute Medicare claims are representative. Therefore, some high volume procedures are represented on the most recent HCFA Prevailing Charge file, but not on The Urban Institute's files from the mid-1970s.

Table 2.7a

RVS Values of High-Volume Medicare Procedures:
Pearson Product-Moment Correlations
(number of observations in parentheses)

	MMA	HIAA	78CAL	74CAL
HCFA	.978	.952	.998	.998
	(95)	(21)	(83)	(82)
MMA		.972	.982	.979
		(21)	(77)	(76)
HIAA			.999	.999
			(13)	(13)
78CAL				.999
				(82)

All correlations significant at .0001 level

Table 2.7b

RVS Values of High-Volume Medicare Procedures:
Spearman Rank-Order Correlations

	MMA	HIAA	78CAL	74CAL
HCFA	.948	.978	.979	.975
	(95)	(21)	(83)	(82)
MMA		.963	.960	.965
		(21)	(77)	(76)
HIAA			.994	.996
			(13)	(13)
78CAL				.994
				(82)

All correlations significant at .0001 level.

The 1974 and 1978 Urban Institute California files enabled us to investigate the effects of time upon charge-based RVSs. By examining the Pearson and Spearman correlations of procedures' values on the two California scales--78CAL and 74CAL--we can ascertain the stability of RVS values for Medicare procedures over a reasonable time interval. As the correlations reveal, procedures' relative values and rankings on the two scales were highly correlated. While we cannot infer a general principal from this single pair of correlations, these values at least do not refute the proposition that charge-based RVSs are robust with respect to small (four years or less) temporal differences in the underlying data.*

Although all the correlations of Table 2.7 were large by conventional standards, some sets of scales were associated with higher r and r_s values than were others. Perhaps most noteworthy is that procedures' values on the MMA-RVS were somewhat less highly correlated with values on the other scales than was the norm in this analysis. One might attribute this finding to the fact that the other four RVSs were constructed from explicit charge data, while MMA-RVS was not.** However, the absolute correlations between scores on MMA and the other RVSs were still well in excess of .90. Therefore, an in-depth investigation into the cause of these somewhat lower correlations is unwarranted.

---

*A reasonable hypothesis is that such high correlations are the result of less-than-rapid changes in the technologies of the procedures under study. It is not clear, however, that such an hypothesis would be valid for other time periods or with respect to other procedures.

**Unit values of procedures on the MMA-RVS were, in part, derived from the expert judgments of panels of physicians who were asked to evaluate the relative values of procedures commonly performed by members of their own medical specialty. Therefore, if such judgments differ from the markets' relative prices, discrepancies could emerge between the MMA-RVS and charge-based scales.

The correlations of Table 2.7 support the proposition that the relative values of high-volume Medicare procedures are not appreciably affected by geographic or temporal differences in the data underlying the RV scales. This proposition applies to the set of Medicare procedures analyzed in toto but may not hold for particular procedures. Indeed, it is possible that the relative values of some procedures are more stable across the different scales than are the scores of others. To investigate this possibility, we analyzed RVS scores of procedures in particular service categories--medicine, surgery, radiology, and pathology. We report product-moment correlations of these scores in Table 2.8 and rank-order correlations in Table 2.9.

The data of Table 2.8 confirm the robustness of relative values of procedures (in specific service categories) across RVSs constructed from alternative data bases. All correlations are greater than .85 in value; squared correlations imply that more than 70 percent of the variation in procedures' scores on any one RVS are explained by variation on the other scales. Inspection of Table 2.8 also reveals that cross-scale correlations are greater for procedures in certain of the service categories. In particular, Pearson correlations of surgical procedures' scale values are above .95 in all cases. In contrast, medical procedures' scores on pairs of RVSs are correlated at the .90 level or greater. Except for the somewhat lower r values between scores on the HCFA and MMA scales, pathology procedures' values on pairs of RVSs were correlated above the .94 level. Of the four procedure categories, radiology scores exhibited the weakest relationship across the five RVSs, Pearson correlations being on the order of .86 or greater. This observation might lead one to conclude that relative values of radiology procedures are the most sensitive to variation in the underlying sources of the scales' data. However, these lower r values might be partially explained by the small number of pathology procedures in the sample frame.

Table 2.8

RVS Values of High Volume Medicare Procedures by Category
Pearson Product-Moment Correlations
(number of observations in parentheses)

Medicine

	MMA	78CAL	74CAL
HCFA	.988 (39)	.964 (38)	.950 (38)
MMA		.899 (35)	.914 (35)
78CAL			.989 (38)

Radiology

	MMA	78CAL	74CAL
HCFA	.981 (9)	.920** (9)	.908** (9)
MMA		.859* (7)	.874* (7)
78CAL			.989 (9)

Surgery

	MMA	HIAA	78CAL	74CAL
HCFA	.968 (26)	.952 (21)	.997 (15)	.997 (15)
MMA		.972 (21)	.968 (15)	.963 (15)
HIAA			.999 (13)	.999 (13)
78CAL				.998 (15)

Pathology

	MMA	78CAL	74CAL
HCFA	.865 (21)	.988 (21)	.977 (20)
MMA		.968 (20)	.943 (19)
78CAL			.986 (20)

*significant at .05 level

**significant at .001 level

all else significant at .0001 level

Table 2.9

RVS Values of High Volume Medicare Procedures by Category
Spearman Rank-Order Correlations
(number of observations in parentheses)

Medicine

	MMA	78CAL	74CAL
HCFA	.850 (39)	.924 (38)	.932 (38)
MMA		.835 (35)	.870 (35)
78CAL			.991 (38)

Radiology

	MMA	78CAL	74CAL
HCFA	.891 (9)	.937** (9)	.971 (9)
MMA		.883* (7)	.883* (7)
78CAL			.979 (9)

Surgery

	MMA	HIAA	78CAL	74CAL
HCFA	.965 (26)	.978 (21)	.992 (15)	.979 (15)
MMA		.963 (21)	.972 (15)	.976 (15)
HIAA			.994 (13)	.996 (13)
78CAL				.988 (15)

Pathology

	MMA	78CAL	74CAL
HCFA	.821 (21)	.919 (21)	.906 (20)
MMA		.866 (20)	.834 (19)
78CAL			.958 (20)

*significant at .05 level

**significant at .001 level

all else significant at .0001 level

In fact, both the statistical significance and magnitude of the Pearson correlations of pathology scores on the different RVSs are likely to be sensitive to the limited sample of pathology procedures available for analysis.

The data in Table 2.8 reveal another interesting pattern. In every case, a correlation coefficient below .90 was associated with the MMA-RVS. Only among surgical procedures were MMA-RVS values correlated at levels of .90 or better with scores on the other four scales. This observation suggests that RVSs are sensitive to adjustments made by experts to align the scales with their beliefs regarding the relative worth of procedures.

The rank-order correlations in Table 2.9 tell essentially the same story as the r values in Table 2.9. Surgical procedures' rankings along all five scales are highly correlated, r_s values ranging from .96 to .99. Medical, radiological, and pathology procedures' rankings are also highly correlated between pairs of RVSs constructed explicitly from charge data. These rankings on MMA-RVS, however, are correlated at somewhat lower levels with their counterparts on the former scales.

F. COMPARISON OF CHARGE-BASED SCALES: SUMMARY

In this chapter we assessed the robustness of charge-based RVSs with respect to variations in the characteristics of the underlying data. We were particularly interested in the consistency of the scores and rankings of high-volume Medicare procedures across alternative charge-based scales. Our analyses revealed that varying the type of charge selected as "representative" of each procedures' worth did not appreciably affect relative values in general. While the relative values of particular procedures might vary across alternatively constructed scales (such as scales based upon billings versus those based upon reimbursements), the relative values of the high-volume procedures, taken as a class, were remarkably stable.

Another major finding is the stability of procedures' relative values across RVSs constructed from different sources of charge data. We developed RVSs--for high-volume Medicare procedures--from five alternative data sources. Two of these were national in scope, two were from a single large state (at different points in time), and the fifth was provided by a state physicians' provider group. We found a strong degree of similarity across RVSs constructed from the first four data sources, all explicitly developed from physicians' charges. The state physicians' association scale (which incorporated substantial judgmental information on the worth of procedures) was somewhat less highly correlated with the other scales.* However, even these smaller correlations were large by conventional standards.

We also investigated the stability of charge-based RVSs over a four-year time period. Here, too, we found consistency as measured by large product-moment and rank-order correlations between procedures' values on the temporally distinct scales.

Finally, we investigated the cross-scale stability of the relative values of procedures in different service categories. On balance, we found that RVS values for procedures in each category were robust with respect to the source of the underlying charge data. While this held more strongly for some procedure groups (notably surgical procedures) than for others, inter-group differences were essentially inconsequential.

---

*We also contrasted the scores of the 103 HCFA procedures on alternatively constructed RVSs--based upon the HCFA Prevailing Charge file--with their scores on the California Relative Value Studies scale. The results were analogous to those reported here vis-a-vis MMA. The correlations between procedures' values on the 1969 CRVS scale and their values on HCFA data-based scales were smaller than the correlations between HCFA scales. Even so, these CRVS vs. HCFA correlations were large by conventional standards--with Pearson values on the order of .90 and above.

CHAPTER 3

THE STATISTICAL COST FUNCTION APPROACH

by

Katherine Swartz

One of the tenets of micro-economic theory is that resources are allocated efficiently if the price of every product equals its marginal cost of production as well as its average cost of production, and the marginal revenue and average revenue it generates. In a world with perfect competition and complete information, market forces generally produce these conditions and an efficient allocation of resources exists. However, when markets are not competitive--perhaps because there is monopolistic competition--or when information is imperfect or when several prices for the same product exist because of insurance, our ideal world of economic models disintegrates. Firms then combine resources inefficiently, since prices of goods exceed their marginal costs of production. In this non-ideal world, basing relative values on prices may not provide us with a system that accurately reflects the relative costs of producing various outputs efficiently.

Another tenet of micro-economic theory provides us with an approach for determining the marginal cost of a product in a non-ideal world. This approach rests on "duality theory," which simply states that under certain conditions, every production function has an underlying, or dual, cost function. And when a production function has a dual cost function, we can estimate the cost function parameters and obtain the marginal cost of producing another unit of a firm's product(s). If we create a relative value scale based on the marginal costs of the different products, we avoid the problem that market prices do not equal marginal costs. In fact, if we create a relative value system that sets the price of each unit of output equal to its marginal cost, we would have both absolute price levels and relative prices for each output. Since this is among the objectives of an ideal relative value scale, it is important to assess the feasibility of estimating cost functions for physicians' practices.

This chapter explores the feasibility of using cost functions for an RVS. It provides some background to the use of cost functions and discusses the particular functional form for cost functions that seems most promising for estimating the marginal costs of producing the various outputs in physicians' offices. The second half of the chapter then turns to the practical considerations of using this approach for constructing an RVS. The first part of the chapter is a very technical discussion; those who do not wish to read it are encouraged to skip ahead to the discussion of the feasibility of estimating a cost function for purposes of constructing an RVS.

A. ECONOMIC THEORY BEHIND THE COST FUNCTION APPROACH

The marginal cost of each of a firm's outputs can be obtained when we know a firm's production process, but in order to estimate a firm's production function we need to know the levels of the inputs used in the production process. Usually, it is very difficult to obtain data on the levels of the inputs used, and this has hindered attempts to estimate production functions. However, "duality theory" states that under certain conditions every production function has an underlying cost function, that is, its "dual," and vice versa. This suggests that, if we can obtain data on input prices, then we can estimate the relevant structural cost function. It is then a straightforward, although mathematically cumbersome, matter of computing the marginal cost of each output. To do this, the cost elasticity for each output is estimated first. The cost elasticity is then multiplied by the predicted value for the average cost of the particular output.*

Until recently, a major problem with estimating cost functions for physicians' practices was the multiproduct nature of the physician firm. Physicians clearly produce many different services. Combining these services into one product for the purpose of statistical estimation is not satisfactory for the purpose of constructing a relative value scale. Over the last five years, however, theoretical work by Greene and Christensen (1976) and Brown, Caves and Christensen (1979) has led to the development of a fairly general, multiproduct cost function that can be estimated.

What Properties Should the Cost Function Have?

If we are going to be able to use this cost function approach for determining an RVS, the cost function ought to have five properties. First, it should have a flexible functional form, by which we mean that we do not want to impose *a priori* conditions on its first and second order derivatives. This is particularly important for looking at physicians' practices, since they obviously can produce more than one type of output (e.g., office visit versus hospital visit versus lab work). The final three properties have more to do with the empirical problems involved with using cost functions. The flexible

---

*For the translog multi-product cost function, the cost elasticities are:

$$\frac{\partial \ln C}{\partial \ln Y_i} = \alpha_i + \sum_{j=1}^{m} \delta_{ij} \ln Y_j + \sum_{j=1}^{n} \varphi_{ij} \ln P_j$$

where there are i = 1, ..., m outputs Y, and j = 1, ..., n inputs prices P. To obtain the marginal cost of each output i:

$$\frac{\partial C}{\partial Y_i} = \left[\alpha_i + \sum_{j=1}^{m} \delta_{ij} \ln Y_j + \sum_{j=1}^{n} \varphi_{ij} \ln P_j\right] \cdot \frac{\hat{C}}{Y_i} = \frac{\partial \ln C}{\partial \ln Y_i} \cdot \frac{\hat{C}}{Y_i}$$

where $\hat{C}$ is the predicted value of the cost function.

form, multi-product cost function should be linearly homogeneous in input prices for all possible price and output levels (which means that a proportionate increase in all inputs leads to the same proportionate increase in the outputs, regardless of the initial levels of the inputs). This condition assures the existence of a dual relationship between the cost and production functions. Fourth, there should be a "small" number of parameters to be estimated. Otherwise, the estimation procedure will be extremely expensive. Fifth, the cost function should allow for zero-valued output quantity levels. This is a non-trivial issue, since many firms may not produce all of the possible outputs (not all surgeons perform all types of operations), and we want to be able to consider them when estimating the cost function.

There have been several flexible functional forms that have been proposed over the last ten years, but only the generalized translog function has all of the desired properties. The translog function uses the natural logarithm for total cost, input prices, and output quantities, but the natural log of zero is not defined, so this will not work for zero output quantities. However, we can use the generalized translog function which uses the Box-Cox metric* only for the output and maintains the natural logarithm for total cost and input prices. The generalized function has only one more parameter than the translog to be estimated, and using the Box-Cox metric for output quantities does not violate the homogeneity requirements.

---

*The Box-Cox metric is: $f_i(Y_i) = (Y_i^{\lambda} - 1)/$ for $\lambda \neq 0$ and $f_i(Y_i) = \ln Y_i$ for $\lambda = 0$. So long as λ is strictly positive, the Box-Cox metric is well-defined for zero output levels: $f_i(0) = -1/\lambda$.

B. FEASIBILITY OF USING THE COST FUNCTION APPROACH

While the theoretical rationale for using the cost function approach to setting up an RVS is relatively straightforward, there are at least three sets of issues related to its implementation that need to be addressed. These issues are discussed in turn below, but in sum they make the prospect for implementing the cost function approach very dubious.

The first implementation issue consists of problems with the numbers of outputs and input prices. Current procedure coding schemes include approximately 5,000 to 6,000 separate procedures, each of which would presumably require a relative value. This volume is simply too large to use in a multi-product cost function. Even if we were to concentrate only on various types of visits (which account for between 60 to 75 percent of revenues for most specialties), numbers of radiology and pathology procedures, and number of operations, we would still be facing more than a dozen types of outputs. Multi-product cost functions with more than two separate outputs have not been estimated to our knowledge.* Thus, the robustness and statistical reliability of parameter estimates from a cost function with this many outputs are not known. Furthermore, it is not clear what happens to the parameter estimates when only a subset of all the possible outputs are included. This issue may be particularly crucial here where there is such a large number of procedures to begin with. Finally, in connection with this area, we do not know what the maximum numbers of outputs and input prices are before the estimation procedure becomes unwieldly or breaks down entirely.

---

*Christensen, Christensen and Schoech (1982) have used five outputs combined via the Tornqvist index number procedure to obtain one aggregate output.

The second set of issues concerns the data used for the outputs and the input prices. The output levels are chosen by the physician, who presumably understands that his costs and output levels are interrelated. Thus, the output levels are endogenous variables, which will cause ordinary least squares estimates of the cost function parameters to be biased. If a simultaneous equations estimation method can be used to provide us with $\hat{Y}_i$, which are estimated from instrumental variables that are uncorrelated with the cost functions' error terms, then this problem can be overcome. However, it is not at all clear that this is possible.

On the input price side, we do not know the prices for all the inputs. In particular, we do not observe the physician's own-time price. Many surveys obtain physicians' net incomes, but we know that net income should include at least the financial returns from being a manager and owner of capital in the physician's office. These net income figures are subject to measurement error also, since doctors (like most small business owners who earn relatively high incomes) are able to defer income.* Thus, it is difficult to obtain good data on a physician's hourly wage for "pure physician" work. It might be possible to impute estimates of such an hourly wage based on data for physicians who are employees in small group practices or HMOs. However, employee-physicians are only a very small proportion of all physicians; and the argument can be made that employee-physicians and physicians who own their practices differ in other characteristics, some of which may be related to earnings, too. Thus, imputing hourly wages for all physicians on the basis of employee-physicians' wages would be subject to sample selection bias and therefore unreliable. A

---

*Since this willingness to defer some income occurs more often as income increases, the measurement error is compounded by heteroskedasticity. The heteroskedasticity problem can be solved by taking the natural log of income, however.

more satisfactory solution to the problem of estimating an hourly wage for "pure physician" work has to be found.

Collecting data on wages for non-physician personnel in a medical practice, and on the prices doctors pay for materials and other nondurable items, should be straightforward. It might be possible to collect these data from sources other than physicians, such as the Bureau of Labor Statistics and the Department of Commerce. But determining annual expenditures for investments in capital equipment and office space is likely to be very difficult--even with the cooperation of physicians. A questionnaire requesting detailed data on tangible assets would have to be quite explicit in questions about rates of replacement and about how depreciation and other accounting procedures are handled. It is likely that a physician's accountant or business manager would have to supply the answers to such questions.

On the brighter side of these data problems, if all these data can be obtained, it might be possible to construct separate Tornqvist indices for non-physician labor, capital, and material.* Constructing such price indices for the three aggregate inputs is not trivial. But the advantage from doing so is that it greatly reduces the number of input prices and, hence, share equations and parameters that need to be estimated.

A final problem that arises within these data issues concerns the mixing of physician specialties. The problem is whether to consider all specialties together or separately. Clearly some procedures are only done by a few specialties, while others can be done by all doctors; but, for example, it may be that procedures that can be done by all doctors will show higher marginal costs of capital when capital-intensive specialties are included as compared

---

*Christensen has done this extensively. See Christensen, Christensen, and Schoech (1982) and Caves, Christensen, and Swanson (1981) for examples.

to when they are not. If these capital-intensive specialties rarely produce such outputs, should they be included? Sensitivity tests need to be performed to determine how the mix of specialties used in the data base affects the parameter estimates.

The final set of issues that arise in implementing an RVS based on the cost function approach concerns its acceptance by non-economists, particularly physicians. Once the methodology has been set up, it will not need a large number of economists or statisticians to continue running it to obtain updated RVSs. But the technique is not easy to understand and it may seem like a lot of "black box" hand waving. Physicians, health insurance executives and others are understandably likely to object to the lack of simplicity in this method. The simplicity critique is compounded by the other implementation issues discussed above. Physicians are particularly likely to object to the fact that their own-time wages are not known and that not all the possible outputs could be considered in the cost function. From just a public acceptance point of view, the cost function approach to constructing an RVS seems impractical.

C. IMPLEMENTATION ASSESSMENT USING PERFORMANCE CRITERIA

To summarize these implementation problems and to compare them with those of other methods for constructing RVSs, we assess the cost function approach using our eight previously identified performance criteria.

Data Requirements

Output quantities are endogenous, so we would have to estimate output quantities based on variables that are not correlated with the error terms of the cost functions. This will be a difficult task. Wages for physicians' own-time will have to be imputed--again a difficult task. Finally, data on capital input prices have to be collected, and this, too, is likely to be a complicated task. Most of this information can only be collected from surveys of physicians.

Technical Expertise Required

Implementing a cost-function approach will require substantial computing capacity, and personnel with advanced training in econometrics and accounting will be needed in the initial set-up stage. But once the technique is in place, it should not be necessary to have large numbers of such personnel involved in the updating of the scale.

Pricing Problem Medical Procedures

The cost function approach is likely to yield unbiased estimates of the marginal costs of doing new or complex procedures if the data problems connected to physicians' own-time wages and capital input prices are satisfactorily resolved. This "if" is unlikely to be satisfied, and hence it is difficult to state a priori how the marginal costs will be biased.

Potential for Future Updating of Scales

Updating a cost function-based RVS would require data on output quantities and input prices and then a re-estimation of the system of equations in order to obtain new parameter estimates. Since the data collection would be costly, updating the RVS is not a simple matter.

Development of Specialty- and Location-Specific Scales

Estimating cost functions for specialty groups and location-specific physicians should be quite simple if enough data are available. The cost function parameters would simply be estimated for the desired subsets of physicians.

Implementation by Individual Insurance Carriers

While individual insurance carriers could each implement a cost function-based RVS, there would be enormous duplication of effort if this should occur. From the standpoint of efficiency, it makes more sense for one RVS to be constructed, perhaps by a consortium of carriers. Carriers are very unlikely to have personnel with the required statistical and computing skills.

Suitability Across Procedures

Since it is not feasible to estimate multi-product cost functions with very many outputs, it appears that this approach is only suitable for a very few procedures. Such procedures are likely to be those that occur most frequently.

External Costs

The cost function-based RVS does not consider costs outside the physician's practice, since they are not part of the inputs. This approach is strictly limited to the practice's marginal costs for each output. The considerations apropos to this criterion raised in the charge-based method discussion (Chapter 2) are relevant here too.

D. FUTURE ANALYSIS

Our analysis of the cost function approach leads us to believe that it is not a promising path to follow for constructing an easily implementable RVS for physicians' fees. However, if the Health Care Financing Administration wishes to pursue the cost function approach on an exploratory basis, we will do so in the project's second year.

CHAPTER 4

TIME-BASED METHODS

by

Jack Hadley

Constructing relative values for physicians' services from the amount of time a physician spends performing a procedure has considerable intuitive appeal. The physician's time is an obviously important component of the cost of providing physicians' services. Time is relatively easy to measure, and the concept of using relative time to construct relative values is easy to understand. The time spent performing a procedure is unlikely to be directly affected by general price inflation or cost-of-living differences. In other words, time may be a more stable yardstick for constructing relative values than monetary values.

From a theoretical vantage point, a time-based relative value scale is a subset of the more general cost-based approaches discussed in chapters 3 (statistical cost functions) and 5 (micro-costing). Physicians' time is one of several inputs used in producing physicians' services. Other inputs include physicians' office employees (nurses, receptionists, technicians), office space, and medical equipment and supplies. In addition, hospitals, clinics, and other institutional providers often contribute inputs, frequently paid for by the patient directly, to the production of physicians' services.

Constructing relative values as the ratios of representative amounts of physicians' time used to perform the services identified by CPT-4 would be a good approximation to cost-based relative values <u>if other inputs are combined with physicians' time in roughly the same proportions across all physicians' services</u>. Thus, a procedure which uses twice as much of a physician's time as another could be reliably ranked as being twice as costly.

Is this a realistic assumption? Within particular classes of procedures, it may be. For example, the full cost of an office visit with a physician is likely to be highly correlated with the amount of time spent with the physician, as may be the benefit or value to the patient. For surgical procedures, the major cost to the physician is her/his own time, since the physician does not pay for most of the other inputs, i.e., surgical assistants, operating room space and equipment, x-ray and post-operative monitoring equipment.

For radiology procedures, however, there may be very little correlation between physicians' time and procedures' values because of major differences among procedures in the costs of the equipment used and in whether the physician is needed to perform the procedure as well as interpret its results. For example, the various types of CT-scanning use equipment which vary significantly in their costs, but the amount of physician time needed for interpretation may vary much less. In fact, if the more expensive equipment produces better quality images, then physicians' time could be inversely correlated with total expense. Furthermore, the tests themselves can usually be performed by technicians. In contrast, other procedures, like the G.I.-series for example, are typically performed by the physician, but use much less expensive equipment than CT-scanning, and often require less extensive training to interpret.

The fixed-input-proportions assumption is even less likely to be tenable across procedure types. Not only do the types and costs of equipment and support staff vary, but the physician's own investment in professional training varies significantly among the specialties associated with different procedures. For example, family practitioners, general internists, and pediatricians typically have three years of residency training compared to four for pathologists, five for general surgeons, and six to eight for surgical subspecialties. These differences in training investment would cause the value of time to vary among specialties. Thus, the basic unit of measure for constructing time-based relative values is in reality a variable, not a fixed yardstick.

The variable yardstick problem is correctable, although any adjustment is necesssarily arbitrary. For example, Hsiao and Stason (1979) computed an "opportunity cost multiplier" in an attempt to adjust for differences in length of training. They assumed that the rate of return to training should be the same for all physicians, regardless of the number of years of training. This means that the value of time in practice has to be higher for specialties with longer training periods, by an amount sufficient to just offset earnings lost by choosing an extra year of training over entering practice. While all of their assumptions, including those about age-earnings profiles, annual hours worked, and discount rates, may be perfectly reasonable, there is nothing necessary or compelling about them. Other, equally reasonable assumptions would lead to different adjustment factors.

Before turning to the evaluation of the time-based methods using the eight performance criteria identified earlier, it is important to point out the distinction between constructing relative values for physicians' services on the basis of time and paying directly for physicians' time. The former uses a designated procedure coding terminology, in this case CPT-4, for identifying the services to be paid. A time-based relative value scale uses some measure of procedure-specific physicians' time spent to construct relative values <u>among procedures</u>. Paying physicians directly on the basis of time implies that the physician would receive a fixed amount of remuneration per unit of time, <u>regardless of</u> the procedures performed. In other words, time units would be the <u>de facto</u> output measures, not procedures as defined by CPT-4.

This study examines only relative value scales for procedures defined by CPT-4. Thus, we do not address the issue of whether differences across physicians or medical specialties in remuneration per unit of time are too high or too low. A time-based relative value scale for CPT-4 procedures will result in a particular relative value for each procedure. Whether some procedures are undervalued or overvalued when relative values are based on time (rather than charges, average total costs, marginal costs, or some other measure of value) can only be assessed by bringing outside information or social preferences to bear. We can examine whether a time-based scale differs from a charge-based scale (and do so below), but contrasting methods does not in and of itself indicate which result is better or more appropriate. These are normative and political decisions which cannot be resolved on technical grounds.

The distinction between a time-based relative value scale for physicans' services and paying for time directly is important for addressing a common criticism of time-based payment: it penalizes the skilled or efficient physician who works more quickly than average (Harris, 1977). This argument is valid for a system that pays physicians directly for their time. It has little relevance for a time-based relative value scale for physicians' services. Physicians who are more skilled or better organized probably produce all services more quickly (i.e., with less direct input of their own time) than the average physician. However, it does not follow that their relative time inputs among procedures will vary significantly compared to a less skilled or less efficient physician. For example, even if they used half as much physician time per procedure as the average physician, their relative time input among procedures would be the same. Similarly, physicians who perform only complicated or difficult services which require substantial time input will not be penalized, since these services would have high time-based relative values.

In the next section of this chapter we evaluate time-based methods using the eight performance criteria identified earlier. We then examine the actual construction of time-based relative values using data from the USC-Mendenhall study of physicians' practices. Most of the analysis focuses on relative values for several of the office, home, and hospital visits identified by CPT-4. As will be discussed more fully, this is primarily because information from the available data bases which report physicians' time cannot be readily converted to the CPT-4 procedural terminology. Although office, home, and hospital visits are only a small proportion of CPT-4 procedure codes, they accounted for over 80 percent of Medicare's expenditures for the 30 procedures most frequently performed by approximately 5,000 California physicians in 1978.* Limited information is available for a few surgical and other procedures. No physicians' time information is currently available for radiology and laboratory procedures.**

---

*Unpublished data, The Urban Institute's California Physicians Medicare Claims File. The most frequently performed procedure was the limited office visit, established patient (CPT-4 code 90050), which occurred 141,083 times in the first calendar quarter of 1978. The 30th most frequently performed procedure was the limited home visit, established patient (CPT-4 code 90150), which occurred 3,196 times.

**The College of American Pathologists has developed data on <u>nonphysician</u> laboratory time requirements for specific laboratory procedures using specific types of equipment. See Chapter 5 below for more information about the <u>Laboratory Workload Recording Method</u> manual.

A. Performance Evaluation of Time-Based RV Scales

Regardless of the general characteristics of a time-based RV scale, it is also important to evaluate this method in terms of its implementation and management problems. This section assesses time-based methods using eight performance criteria.

Data Requirements

In order to implement a time-based RV scale for all physicians' services, one would need a large quantity of data on physicians' time spent for each of the procedures to be included in the scale. No currently available data base contains the requisite data for more than a few procedures. As described more fully in another report (Carlson, 1982), the most recent data on physicians' time are collected by the National Ambulatory Medical Care Survey (NAMC), an annual survey, conducted by the National Center for Health Statistics, of a national probability sample of about 2,000 physicians' offices.

There are three major drawbacks to using this data to construct time-based RV scales. First, no procedure codes are used to identify the services provided during the visit. The survey records symptoms and reasons for visits, which cannot be translated into an existing procedure coding terminology. Second, the survey is limited to visits which occur in physicians' offices. Third, the few therapeutic and diagnostic services identified by the form are checked off if the service is either performed or ordered. Figure 4.1 reproduces the typical patient record.

The other major data base containing information on physicians' time was constructed by the USC-Mendenhall study of physicians' practices. Using log diaries, the project surveyed approximately 10,000 physicians (resulting in about 250,000 patient-physician encounters) between 1974 and 1976. Like the NAMC survey, services are identified mainly by diagnosis and etiology, rather

Figure 4.1

NAMC Patient Record

1979 Patient Record

ASSURANCE OF CONFIDENTIALITY — [illegible]

Department of Health, Education, and Welfare
Public Health Service
Office of Health Research, Statistics, and Technology
National Center for Health Statistics

B

1. DATE OF VISIT ___/___/___ (Mo. Day Yr.)

PATIENT RECORD
NATIONAL AMBULATORY MEDICAL CARE SURVEY

TIME OF VISIT

a.m. / p.m.

a.m. / p.m.

2. DATE OF BIRTH ___/___/___ (Mo. Day Yr.)

3. SEX
1 ☐ FEMALE
2 ☐ MALE

4. COLOR OR RACE
1 ☐ WHITE
2 ☐ BLACK
3 ☐ ASIAN, PACIFIC ISLANDER
4 ☐ AMERICAN INDIAN/ ALASKAN NATIVE

5. ETHNICITY
1 ☐ HISPANIC ORIGIN
2 ☐ NOT HISPANIC

6. WAS PATIENT REFERRED FOR THIS VISIT BY ANOTHER PHYSICIAN?
1 ☐ YES
2 ☐ NO

7. PATIENT'S COMPLAINT(S), SYMPTOM(S), OR OTHER REASON(S) FOR THIS VISIT (In patient's own words)
a. MOST IMPORTANT ______
b. OTHER ______

8. MAJOR REASON FOR THIS VISIT (Check one)
1 ☐ ACUTE PROBLEM
2 ☐ CHRONIC PROBLEM, ROUTINE
3 ☐ CHRONIC PROBLEM, FLAREUP
4 ☐ POST SURGERY/ POST INJURY
5 ☐ NON ILLNESS CARE (ROUTINE PRENATAL, GENERAL EXAM, WELL BABY, ETC.)

9. PHYSICIAN'S DIAGNOSES
a. PRINCIPAL DIAGNOSIS/PROBLEM ASSOCIATED WITH ITEM 7a ______
b. OTHER SIGNIFICANT CURRENT DIAGNOSES ______

10. HAVE YOU SEEN PATIENT BEFORE?
1 ☐ YES 2 ☐ NO
IF YES, FOR THE CONDITION IN ITEM 9a?
1 ☐ YES 2 ☐ NO

11. DIAGNOSTIC SERVICES THIS VISIT (Check all ordered or provided)
1 ☐ NONE
2 ☐ LIMITED HISTORY/EXAM
3 ☐ GENERAL HISTORY/EXAM
4 ☐ PAP TEST
5 ☐ CLINICAL LAB TEST
6 ☐ X-RAY
7 ☐ BLOOD PRESSURE CHECK
8 ☐ EKG
9 ☐ VISION TEST
10 ☐ ENDOSCOPY
11 ☐ MENTAL STATUS EXAM
12 ☐ OTHER (Specify) ______

12. THERAPEUTIC SERVICES THIS VISIT (Check all ordered or provided)
1 ☐ NONE
2 ☐ DRUG (PRESCRIPTION)
3 ☐ DRUG (NONPRESCRIPTION)
4 ☐ INJECTION
5 ☐ IMMUNIZATION/ DESENSITIZATION
6 ☐ DIET COUNSELING
7 ☐ FAMILY PLANNING
8 ☐ MEDICAL COUNSELING
9 ☐ PHYSIOTHERAPY
10 ☐ OFFICE SURGERY
11 ☐ PSYCHOTHERAPY/ THERAPEUTIC LISTENING
12 ☐ OTHER (Specify) ______

13. DISPOSITION THIS VISIT (Check all that apply)
1 ☐ NO FOLLOW-UP PLANNED
2 ☐ RETURN AT SPECIFIED TIME
3 ☐ RETURN IF NEEDED, P.R.N.
4 ☐ TELEPHONE FOLLOW-UP PLANNED
5 ☐ REFERRED TO OTHER PHYSICIAN
6 ☐ RETURNED TO REFERRING PHYSICIAN
7 ☐ ADMIT TO HOSPITAL
8 ☐ OTHER (Specify) ______

14. DURATION OF THIS VISIT (Time actually spent with physician)
______ MINUTES

15. IS THE REASON FOR THIS VISIT? (Check one)
1 ☐ ACCIDENTAL INJURY (Answer 16-19)
2 ☐ PRODUCT RELATED ILLNESS (Answer 16-19)
3 ☐ NEITHER OF THE ABOVE → STOP (Go to next patient)

16. DESCRIBE ALL OBJECTS, PRODUCTS, OR SUBSTANCES INVOLVED IN THE ACCIDENT OR PRODUCT RELATED ILLNESS (Be Specific)
______

17. LOCATION OF ACCIDENT OR EXPOSURE TO PRODUCT (Check One)
1 ☐ PRIVATE RESIDENCE
2 ☐ ELSEWHERE (Specify) ______
3 ☐ UNKNOWN

18. WAS PATIENT AT WORK, JOB OR BUSINESS WHEN ACCIDENT OR EXPOSURE OCCURRED?
1 ☐ YES
2 ☐ NO
3 ☐ UNKNOWN

19. WAS PATIENT PREVIOUSLY TREATED FOR THIS CONDITION? (Check all that apply)
1 ☐ NO
2 ☐ YES HOSPITAL EMERGENCY ROOM
3 ☐ YES PRIVATE PHYSICIAN'S OFFICE
4 ☐ YES PHYSICIAN ELSEWHERE (Specify) ______
5 ☐ YES PLACE UNKNOWN
6 ☐ UNKNOWN

than by procedure code. As shown in Figures 4.2 and 4.3, which reproduce a typical encounter form and coding key from the log diary, much more detail was collected than on the NAMC survey. As is described in the methodology section, it is possible to construct CPT-4 procedure codes for a small number of services, mainly visits of various lengths, complexity, and location.

Full scale implementation of a time-based RV scale would require a major new data collection effort to obtain information on the time physicians spend in providing specific services. In order to be useful for reimbursement purposes, services would have to be identified by their CPT-4 (or other acceptable) procedure coding system. The sample size would have to be very large if one wanted to collect data on infrequently performed procedures. As a benchmark for the cost of such a survey, the cost of fielding the NAMC survey in 1981 was approximately $700,000.*

An alternative approach for obtaining estimates of time per procedure, especially for infrequently performed procedures, is to use panels of experts to establish normative time requirements. This would get around the cost of surveying a large number of physicians in order to obtain enough observations to compute reliable estimates of actual time required. (See chapter 6 for the evaluation of the expert panel approach.)

Technical Expertise Required

Survey design and management expertise would be required to collect the necessary data. Once the data are obtained, only routine computing skills and computer capacity would be needed to actually generate the scales.

---

*Personal communication, National Center for Health Statistics.

Figure 4.2

USC-Mendenhall Reporting Form

PLEASE REVIEW INSTRUCTIONS BEFORE STARTING

ENCOUNTER LOCATION			
	1. Office	4. Hospital	7. Industry/School
	2. OPD	5. Emergency Room	8. Home
	3. Clinic	6. Extended Care (ICF, SNF, Nursing Home)	9. Other

PRIMARY PROBLEM TYPE		
	1. Preventive	5. Medical & Surgical
	2. Medical	6. Obstetrical
	3. Surgical - minor	7. Psycho/Social
	4. Surgical - major	8. Environmental/Economic

Tear off here

Staff may complete this section

Use numbers from Coding Key

PATIENT ID

Your Use Only

Line Number	ENCOUNTER DATE and TIME	PATIENT & ENCOUNTER CHARACTERISTICS	PATIENT'S PROBLEM(S)	DIAGNOSTIC PROCEDURES	THERAPEUTIC PROCEDURES	ENCOUNTER CLASS	STATUS
	Mo. Day Yr (Date); Time in Minutes	Location (use one of the above codes 1-9); Age: Estimate to nearest year "0" = under one year; Sex: 1. Male 2. Female; Seen patient before? 1. Yes 2. No; Regular patient? 1. Yes 2. No; Patient Source: 1. Self/Patient 2. Agency 3. Physician 4. Consult; Majority of Care: This patient? 1. Yes 2. No 3. Undetermined; Patient's family? 1. Yes 2. No 3. Undetermined; Your Role: 1. Provide 2. Assist; Number of visits for this problem?; Type (use one of the above codes 1-8)	Primary: Focus, Etiology (Dx); Other: Focus, Etiology (Dx); Focus, Etiology (Dx)	Record all ordered, performed, or Interpreted	Record all ordered or performed	Complexity, Severity, Urgency	Your next contact, Referral/Consultation
101	Primary Dx (Print)		Other Dx (Print)				
102	Primary Dx (Print)		Other Dx (Print)				
103	Primary Dx (Print)		Other Dx (Print)				
104	Primary Dx (Print)		Other Dx (Print)				
105	Primary Dx (Print)		Other Dx (Print)				

Figure 4.3

USC-Mendenhall Coding Key

CODING KEY

FOCUS

1. Well patient
2. Skin, incl. nails and hair
3. Head/Neck
4. Eyes
5. Ears
6. Nose
7. Mouth/Throat
8. Breasts
9. Respiratory - upper
10. Respiratory - lower
11. Circulatory - heart
12. Circulatory - other
13. Blood/Lymphatics
14. Liver/Gallbladder
15. Gastrointestinal - other
16. Arthritis/Rheumatism
17. Muscular - other
18. Skeletal - other
19. Endocrine
20. Neurological
21. Emotional
22. Renal
23. Urinary - other
24. Genital - male
25. Gynecological
26. Obstetrical
27. Multi-system*

*Use only if one of the above is not more descriptive and correct.

ETIOLOGY (DX)

1. None
2. Health assessment
3. Developmental
4. Pre/post-surgical
5. Iatrogenic
6. Idiopathic (based on workup)
7. Infection/post-infection
8. Inflammation
9. Hemorrhaging/Bleeding
10. Occlusion
11. Ulceration
12. Immunologic
13. Allergic
14. Metabolic
15. Deficiencies
16. Nutrition, incl. weight
17. Congenital
18. Degeneration, incl. aging
19. Hypertension
20. Neoplasm - benign
21. Neoplasm - malignant
22. Neoplasm - undetermined

INTOXICATION/ADDICTION

23. Alcohol
24. Drug
25. Poison
26. Tobacco

TRAUMA

27. Burn
28. Concussion
29. Contusion
30. Dislocation/Fracture
31. Foreign body
32. Laceration
33. Sprain

PSYCHO/SOCIAL

34. General adjustment
35. Depression/Anxiety
36. Neuroses
37. Psychoses
38. Psychosomatic

UNDETERMINED

39. Cough
40. Fatigue
41. Fever
42. Headache
43. Malingering
44. Pain

DIAGNOSTIC PROCEDURES

1. None
2. Routine lab: CBC, Urinalysis
3. Serology/VD
4. Blood chemistry
5. "Panel" - automated (e.g., SMA-12)
6. ECG
7. Culture
8. Enzymes
9. Biopsy
10. Bone marrow
11. Endoscopy

RADIOLOGY

12. Chest
13. GI
14. Fluoroscopy
15. Isotopes
16. X-ray - other

OTHER

17. Catheterization
18. EEG
19. Thermography
20. Glucose tolerance
21. Lipid screen
22. Liver function
23. Lumbar puncture
24. Pulmonary function
25. Renal function
28. Thyroid function
27. Breast exam
28. Pap smear
29. Skin test - allergy
30. Skin test - TB/Mycosis
31. Audiometry
32. Tonometry
33. Developmental screen
34. Tests - other

THERAPEUTIC PROCEDURES

1. None
2. Immunizations
3. Injections - other

DISLOCATIONS/FRACTURES

4. Reduction - open
6. Reduction - closed

SURGICAL

6. Minor tissue removal
7. Incision/Drainage
8. Suturing only
9. Debridement
10. Organ removal/repair
11. Foreign body removal
12. Dilation/Curettage
13. Arthrocentesis
14. Paracentesis
16. Catheterization
16. Cauterization/Cryotherapy
17. Surgery - other

POST-SURGERY/INJURY

18. Cast/Splint: apply/remove
19. Dressing: apply/remove
20. Suture removal

OTHER RX

21. Anesthetization
22. Chemotherapy
23. Contraceptives, incl. IUD
24. Physical rehabilitation
25. Radiation therapy
26. Translusion
27. Rx - other

COUNSELING

28. Growth/Development
29. Family/Social/Sexual
30. Patient education
31. Therapeutic listening
32. Treatment program

PRESCRIPTIONS/ORDERS

33. Drugs - systemic
34. Drugs - topical
35. Exercise/Diet
36. Prosthesis/Aids

ENCOUNTER CLASSIFICATION

COMPLEXITY

1. Minimal
2. Brief
3. Limited
4. Extended
5. Comprehensive

SEVERITY

1. None
2. Minor - acute
3. Minor - chronic
4. Moderate - acute
6. Moderate - chronic
6. Severe - acute
7. Severe - chronic

URGENCY

1. None
2. Could have been deferred
3. Needed to see today
4. Should have seen sooner
5. Emergency

STATUS/DISPOSITION

YOUR NEXT CONTACT

1. None this problem
2. Telephone

OUTPATIENT

3. Return as necessary
4. Visit scheduled
6. House call

INPATIENT

6. At hospital admission
7. On rounds (prior admission)
8. In surgery

OTHER

9. Emergency room
10. Extended care (ICF, SNF, Nursing Home)
11. DOA/Death
12. Other

REFERRAL/CONSULTATION

1. None
2. Return to referral source
3. Medical specialist for care
4. Medical specialist for consult
5. Surgical specialist for care
6. Surgical specialist for consult
7. Dentist
8. Physician Extender (e.g., Medex, Nurse Practitioner, Physician's Assistant)
9. Therapist/technician (e.g., Dietary)
10. Public agency
11. Other

Pricing Problem Procedures

A time-based method may be better than a charge-based method for valuing either new procedures subject to the "learning curve" phenomenon or procedures characterized by rapid technical change. If the amount of physicians' time per procedure changes rapidly, then frequent time measurements would be able to capture changes in the procedure's time-based relative value. Charge-based methods may not capture these changes as well because insurance (or some other mechanism) may keep charges high even when the costs of production are falling.

Time-based methods have no particular advantage, however, in valuing infrequently performed procedures. The problem of inadequate charge data would simply be replaced by inadequate time data.

Potential for Future Updating

Each future updating of a time-based RV scale would require a new survey of physicians, since time is not routinely reported as part of any current data collection system which uses a common procedure coding terminology. In addition to the expense of conducting such surveys, it is also possible that physicians' reporting of time spent would be affected by knowing that the survey data would be used to update an RV scale.

Another potential consequence of a time-based RV scale is that it may distort future choices about how to combine inputs to produce medical services. In general, tying the rate of payment for an output or service to the use of a particular input will distort input choice decisions toward using more of the particular input than would be technically efficient. This has been demonstrated for public utility rate setting which sets rates as a function of the cost of capital to the utility--the firm could produce its output at a more efficient capital-labor mix, but to do so would result in lower profits (Averch and Johnson, 1962). Similarly, a physician might be reluctant to adopt technologies or hire aides which substitute for his/her time if such a change would ultimately lower the relative value of a service. Along the same line, physicians might have greater incentives to shift their practices into hospitals or other institutional settings that are reimbursed independently of physicians on a basis other than an RV scale.

Development of Specialty- and Location-Specific Scales

Data availability is the only constraint on constructing specialty- and location-specific scales. If the numbers of physicians surveyed in various specialties or locations were large enough to generate a sufficient number of time observations for relevant procedures, then specialty- and/or location-specific scales would be easy to construct.

Implementation by Local Insurance Carriers

Local carriers could easily construct and implement time-based scales if the required data were available. Carriers would find it extremely difficult and costly to collect such data on their own because of the survey management expertise and large sample sizes that would be required.

Suitability across Procedure Types

As discussed at some length in the introduction of this chapter, time-based RV scales are not very well suited for comparisons across procedure types because of substantial differences in the quantities and mixes of nonphysician inputs needed to produce different types of procedures. Systematic differences across specialties in physicians' length of training also make cross-procedure scales suspect.

External Costs

The amount of time physicians spend producing particular procedures does not reflect or incorporate any costs external to the physician. In fact, as noted in the discussion of updating, there may be incentives to increase total costs by shifting nonphysician inputs to institutional providers or choosing an inefficient mix of physician time relative to other inputs.

B. Methods of Constructing Time-Based RV Scales

Given data on physicians' time and decision on whether and how to group physicians or procedures, the basic method for constructing relative values is essentially the same as the one used to construct charge-based relative values. First, all of the individual time observations for a particular procedure are arranged in ascending order of magnitude. Second, some value from this distribution, for example, the mean, the median (50th pecentile), the 75th percentile, or the 90th pecentile, is chosen to be the representative absolute value for that particular procedure. Third, the absolute value of each procedure is divided by the absolute value of the base or numeraire procedure to compute the relative value for each procedure. (As discussed in chapter 2, the choice of a base procedure is arbitrary and does not alter either the ordinal or cardinal rankings of individual procedures.)

In order to compensate for the nonproportionality of other input costs to physicians' time, one could construct various "multipliers" to adjust physicians' time, up or down, to account for other factors. For example, Hsiao and Stason (1979) state "Not all time is equal; rather the degree of skill and intensity of effort required per unit of time vary widely from one service to another." They identified three factors in addition to time which need to be incorporated into a procedure's relative value: complexity, the opportunity cost of physicians' time (investment in training), and office expenses.

The complexity multiplier was constructed from complexity rankings of the 26 surgical procedures in their study made by 25 board-certified physicians, 5 each from general surgery, obstetrics and gynecology, opthalmology, orthopedics, and urology.

The opportunity cost multiplier was based on the assumption that each specialty should earn the same rate of return on its investment in training. Data were needed for the length of training by specialty, number of years of active practice, residents' salaries by year of training, and specialists' incomes by length of time in practice. They also assumed that variations in hours of work were unimportant, that appropriate discount rates are 7 or 10 percent, that residents' salaries increased at 15 percent per year after the first year of training, and that a general practitioner's average income is the relevant measure of foregone earnings while in training. Another important implicit assumption they made is that each surgical procedure is performed exclusively by a single specialty.

Differences in office expenses were assumed to be related to the proportions of gross receipts going to office expenses for general practitioners, general surgeons, and obstetrician-gynecologists. The average proportion for the two surgical specialties was applied to the other three specialties in their study. They also assumed that these expenses were spread evenly among all procedures produced by each physician.

As can be seen, the process of adjusting physicians' time for other factors involves many arbitrary assumptions. Some of these were necessitated by the lack of data; others are arbitrary no matter what data are available. Obviously, obtaining the additional data needed to compute the various multipliers would add considerably to the expense and complexity of obtaining data on physicians' time.

The time-based RV scales we construct and examine in the next section are based on distributions of unadjusted physicians' time per procedure. Most of the information needed to make various adjustments is not available. Furthermore, we are skeptical that valid and acceptable multipliers could be developed for the purpose of constructing a time-based RV scale to be used in reimbursing physicians.

We will use data from the USC-Mendenhall study of physicians' practices to construct four RV scales for fifteen visit procedures and thirteen other procedures. The data include patient-physician encounters from physicians in five specialties, general and family practice, internal medicine, general surgery, and pediatrics.

The four scales are based on the mean, median, 75th percentile, and 90th percentile of the distribution of physicians' time for each procedure. As discussed above, this data base does not use any standard procedure coding terminology to identify procedures. Information on new or established patient status, encounter location, number of visit, complexity, diagnosis, and etiology was used to construct CPT-4 equivalent procedures. Figures 4.4 and 4.5 identify the CPT-4 procedures and the corresponding USC-Mendenhall information.

After constructing the scales, we compute Pearson and Spearman (rank order) correlation coefficients to determine how sensitive the scales are to the choice of different points from the distribution of physicians' time to be the "representative" absolute value. Separate correlations are also computed for the visit and other procedures to explore the sensitivity of the scales to combining different procedure types.

Figure 4.4

CPT-4 Visit Codes and Corresponding
USC-Mendenhall Visit Codes

Description	CPT-4	USC-Mendenhall Encounter Location Code[3]	Visit Counter[4]	Complexity Code[5]
New Patient, Office Visits[1]				
Brief Evaluation	90000	1-3,5	any	2
Initial Limited	90010	1-3,5	1	3
Initial Comprehensive	90020	1-3,5	1	5
Established Patient, Office Visits[2]				
Minimal	90030	1-3,5	any	1
Brief	90040	1-3,5	any	2
Limited	90050	1-3,5	any	3
Extended re-exam	90070	1-3,5	2+	4
Comprehensive re-exam	90080	1-3,5	2+	5
Established Patient, Home Visit[2]				
Brief	90140	8	any	2
Limited	90150	8	any	3
Hospital Visits				
Initial, Brief or Limited	90200	4	1	2,3
Initial, Comprehensive	90220	4	1	5
Brief Follow-up	90240	4	2+	2
Limited Follow-up	90250	4	2+	3
Extended Follow-up	90270	4	2+	4

Notes:
1. "Seen patient before?" = 2 on Figure 4.2.
2. "Seen patient before?" = 1 on Figure 4.2.
3. See ENCOUNTER LOCATION box on Figure 4.2.
4. See "Number of visits this problem?" on Figure 4.2.
5. See ENCOUNTER CLASS on Figure 4.2 and ENCOUNTER CLASSIFICATION - COMPLEXITY on Figure 4.3.

Figure 4.5

CPT-4 Procedure Codes and Corresponding USC-Mendenhall Identifiers

Description	CPT-4	USC-Mendenhall Procedure[a]
Internal Medicine, General Practice, Family Practice		
Arthrocentesis	20610	T13
Chemotherapy	96030	T22
ECG	93000	D06
Lumbar Puncture	93200	D23
Pediatrics		
Chemotherapy	96030	T23
Surgical		
Arthrocentesis	20610	D35
Catherization (Heart)	93527	D29
Cholecystectomy	47600	T08
Herniorrhaphy	49505	T14
Hysterectomy	58265	T19
I & D	10000	T28
Mastectomy	19160	T06
Proctosigmoidoscopy	45300	D37
Thoracentesis	32000	T17

Note: a. D-Diagnostic, T-Therapeutic. See Figure 4.3 for Coding Key for internal medicine, general practice, and family practice.

C. RESULTS

Table 4.1 reports the absolute and relative values, scale rank, and numbers of observations for the 28 procedures (identified by CPT-4 code and name) in the analysis. The limited office visit, established patient (90050) is the numeraire of each scale. The numbers of observations vary from 29,736 for the numeraire procedure to 8 for the procotosigmoidoscopy (45300) and 13 for heart catherization (93527).

Simple visual examination of the scale values shows that the mean and median scales are more compact, i.e., have smaller means and standard deviations, than the 75th and 90th percentile scales. It appears that this is due mainly to differences in the relative values of the nonvisit procedures, which have the highest relative values. A possible explanation of the greater dispersion for these procedures on the 75th and 90th percentile scales is that these observations are more likely to include procedures actually performed by the physician, while the mean and median scales are mainly for visits during which the physician ordered the procedure. (The data base does not distinguish between ordered and performed.) These procedures also have the lowest frequency of performance. As a result, their distributions may be less reliable estimates of their true time distributions than those of the other procedures.

The scale based on mean time also differs from the other three in that no two procedures have identical relative values. In contrast, the other three scales have many more ties in ranking--the median and 90th percentile scales have only 13 separate values, and the 75th percentile scale has only 16. This may reflect the tendency to report time in relatively large intervals, e.g., 5, 10, or 15 minutes, rather than precise elapsed time.

Table 4.1

Four Time-Based RV Scales, 28 Procedures
(USC-Mendenhall Data Base)

Procedure Name CPT-4 Code (No. of Observations)	Scale Mean	 Median	 75th Percentile	 90th Percentile
All Procedures				
Mean	2.65	2.45	3.24	3.42
Standard Deviation	2.35	2.42	3.70	3.54
1. Brief H.V. Estab. Pat. 90240 (6,406)				
Absolute Value[c]	8.04	5.00	10.00	15.00
Relative Value	.68	.50	.67	.75
Scale Rank	1.00	1.00	2.00	2.50
2. Minimal O.V. Estab. Pat. 90030 (13,103)				
Absolute Value	8.61	7.00	10.00	15.00
Relative Value	.72	.70	.67	.75
Scale Rank	2.00	2.00	2.00	2.50
3. Brief O.V. Estab. Pat. 90040 (26,843)				
Absolute Value	9.13	10.00	10.00	15.00
Relative Value	.77	1.00	.67	.75
Scale Rank	3.00	6.00	2.00	2.50
4. Brief O.V. New Pat. 90000 (4751)				
Absolute Value	10.08	10.00	15.00	15.00
Relative Value	.85	1.00	1.00	.75
Scale Rank	4.00	6.00	6.50	2.50
5. Limited H.V. Estab. Pat. 90250 (7905)				
Absolute Value	11.44	10.00	15.00	20.00
Relative Value	.96	1.00	1.00	1.00
Scale Rank	5.00	6.00	6.50	5.50
6. Limited O.V. Estab. Pat.[a] 90050 (29,736)				
Absolute Value	11.87	10.00	15.00	20.00
Relative Value	1.00	1.00	1.00	1.00
Scale Rank	6.00	6.00	6.50	5.50
7. Brief H.V. New Pat. 90200 (3949)				
Absolute Value	13.13.	10.00	15.00	25.00
Relative Value	1.11	1.00	1.00	1.25
Scale Rank	7.00	6.00	6.50	8.00
8. Chemotherapy 96030 (519)				
Absolute Value	13.88	10.00	15.00	25.00
Relative Value	1.17	1.00	1.00	1.25
Scale Rank	8.00	6.00	6.50	8.00

Table 4.1 (Continued)

Procedure Name CPT-4 Code (No. of Observations)	Scale			
	Mean	Median	75th Percentile	90th Percentile
9. Limited O.V. New Pat. 90010 (4456)				
Absolute Value	14.63	15.00	18.00	25.00
Relative Value	1.23	1.50	1.20	1.25
Scale Rank	9.00	14.00	10.00	8.00
10. ECG 93000 (2264)				
Absolute Value	16.19	12.00	20.00	30.00
Relative Value	1.36	1.20	1.33	1.50
Scale Rank	10.00	10.00	12.00	13.00
11. Extended H.V. Estab. Pat. 90270 (4571)				
Absolute Value	16.67	10.00	15.00	30.00
Relative Value	1.40	1.00	1.00	1.50
Scale Rank	11.00	6.00	6.00	13.00
12. I + D 10000 (138)				
Absolute Value	16.72	15.00	20.00	30.00
Relative Value	1.41	1.50	1.33	1.50
Scale Rank	12.0	14.00	12.00	13.00
13. Extended O.V. Estab. Pat. 90070 (3595)				
Absolute Value	17.85	15.00	21.00	30.00
Relative Value	1.50	1.50	1.40	1.50
Scale Rank	13.0	14.0	13.0	13.0
14. Brief Home V. Estab. Pat. 90140 (53)				
Absolute Value	18.08	15.00	22.50	30.00
Relative Value	1.52	1.50	1.50	1.50
Scale Rank	14.00	14.00	14.00	13.00
15. Lumbar Puncture 93200 (109)				
Absolute Value	19.29	15.00	25.00	40.00
Relative Value	1.62	1.50	1.67	2.00
Scale Rank	15.00	14.00	16.00	16.00
16. Limited Home V. Estab. Pat. 90150 (117)				
Absolute Value	20.04	20.00	25.00	30.00
Relative Value	1.69	2.00	1.67	1.50
Scale Rank	16.00	20.00	16.00	13.00
17. Arthrocentesis 20610 (6,226)				
Absolute Value	24.12	15.00	30.00	58.59
Relative Value	2.03	1.50	2.00	2.93
Scale Rank	17.0	14.0	18.0	18.0
18. Comprehensive O.V. New Pat 90020 (1,485)				
Absolute Value	25.68	20.00	30.00	45.00
Relative Value	2.16	2.00	2.00	2.25
Scale Rank	18.00	20.00	18.00	17.00

Table 4.1 (Continued)

Procedure Name CPT-4 Code (No. of Observations)	Scale Mean	Median	75th Percentile	90th Percentile
19. Comprehensive O.V. Estab. Pat. 90080 (1,246)				
Absolute Value	26.98	20.00	40.00	60.00
Relative Value	2.27	2.00	2.67	3.00
Scale Rank	19.00	20.00	19.00	19.00
20. Comprehensive Hosp. V. New Pat. 90220 (1,088)				
Absolute Value	37.30	20.00	45.00	90.00
Relative Value	3.14	2.00	3.00	4.50
Scale Rank	20.00	20.00	20.00	21.00
21. Herniorrhaphy 49505 (136)				
Absolute Value	41.82	32.50	60.00	90.00
Relative Value	3.52	3.25	4.00	4.50
Scale Rank	21.00	24.00	22.00	21.00
22. Thoracentesis 32000 (55)				
Absolute Value	44.11	30.00	60.00	90.00
Relative Value	3.71	3.00	4.00	4.50
Scale Rank	22.00	23.00	22.00	21.00
23. Hysterectomy 58265 (30)				
Absolute Value	56.13	62.50	90.00	118.50
Relative Value	4.73	6.25	6.00	5.92
Scale Rank	23.00	26.00	24.00	23.00
24. Mastectomy 19160 (46)				
Absolute Value	58.87	17.50	97.50	180.00
Relative Value	4.96	1.75	6.50	9.00
Scale Rank	24.00	17.00	25.00	26.00
25. Cholecystectomy 47600 (82)				
Absolute Value	59.62	60.00	90.00	120.00
Relative Value	5.02	6.00	6.00	6.00
Scale Rank	25.00	25.00	24.00	24.00
26. Proctosigmoidoscopy[b] 45300 (8)				
Absolute Value	62.00	30.00	131.25	180.00
Relative Value	5.22	3.00	8.75	9.00
Scale Rank	26.00	23.00	26.00	26.00
27. Colon Resection 44140 (65)				
Absolute Value	98.25	100.00	177.50	180.00
Relative Value	8.27	10.00	11.80	9.00
Scale Rank	27.00	28.00	27.00	26.00
28. Heart Catherization 93527 (13)				
Absolute Value	120.92	90.00	240.00	312.00
Relative Value	10.20	9.00	16.00	15.60
Scale Rank	28.00	27.00	28.00	28.00

Notes: a. Numeraire procedure.

b. Only includes procedures from general surgery log diaries.

c. Time measured in minutes.

The data on mean time per visit procedure show that the ordering of visits by nominal duration, i.e., minimal, brief, limited, extended, and comprehensive, is consistent with the actual mean times. Within each duration category, visits with new patients take more time than established-patient visits. Home visits take the most time of any duration category, but there is no consistent pattern for office visits compared to hospital visits of a given duration.

Pearson and Spearman correlation coefficients suggest that the four scales are substantially alike (Table 4.2). The Pearson coefficients range from 0.81 to 0.99. The Spearman coefficients are generally higher in value, with a minimum of 0.91, primarily because the greater dispersion of relative values in the 75th and 90th percentile scales does not alter procedures' rank ordering very much. Like the charge-based RV scales constructed in chapter 2, the time-based scales appear to be quite robust with respect to the distribution point chosen to represent each procedure's absolute value. This conclusion must be qualified, however, because information was available only for 28 procedures.

How does a time-based scale compare to a charge-based scale? If the two are very similar, then there would be less justification for undertaking the expensive data collection needed to construct a time-based RV scale. Table 4.3 compares the relative values and ranks of 25 procedures common to the time-based data set and the HCFA prevailing charge file. (Scales based on mean values of time and prevailing charges are used to represent the two families of scales.) Not surprisingly, there are substantial differences in the relative values assigned to the nonvisit procedures. Grouping procedures into surgery (herniorraphy, hysterectomy, cholecystectomy, colon resection and heart catherization), hospital visits, office visits, and all other procedures

Table 4.2

Pearson and Spearman Correlation
Coefficients, Time-Based RV Scales, 28 Procedures
(USC-Mendenhall Data Base)

	Pearson Correlation Coefficients			
	Mean	Median	75th Percentile	90th Percentile
Mean	1.00	.92	.99	.97
Median		1.00	.90	.81
75th Percentile			1.00	.98
90th Percentile				1.00

	Spearman Correlation Coefficients			
	Mean	Median	75th Percentile	90th Percentile
Mean	1.00	.95	.99	.99
Median		1.00	.95	.91
75th Percentile			1.00	.98
90th Percentile				1.00

Table 4.3
Time-Based and Charge-Based
RV Scales, 25 Procedures

Procedures (CPT-4 Code)	Absolute Value Relative Value (Scale Rank) Time-Based[a]	 Charge-Based[b]
1. Brief H.V. Estab. Pat. (90240)	8.04 .68 (1.00)	15.52 .96 (3.00)
2. Minimal O.V. Estab. Pat. (90030)	8.61 .72 (2.00)	8.26 .51 (1.00)
3. Brief O.V. Estab. Pat. (90040)	9.13 .77 (3.00)	13.72 .85 (2.00)
4. Brief O.V. New Pat. (90000)	10.08 .85 (4.00)	21.34 1.32 (8.00)
5. Limited H.V. Estab. Pat. (90250)	11.44 .96 (5.00)	20.11 1.25 (5.00)
6. Limited O.V. Estab. Pat. (90050)	11.87 1.00 (6.00)	16.10 1.00 (4.00)
7. Brief H.V. New Pat. (90200)	13.13 1.11 (7.00)	39.69 2.46 15.00)
8. Chemotherapy (96030)	13.88 1.17 (8.00)	20.91 1.30 (6.00)
9. Limited O.V. New Pat. (90010)	14.63 1.23 (9.00)	27.38 1.70 (12.00)
10. ECG (93000)	16.19 1.36 (10.00)	27.68 1.71 (13.00)
11. Extended H.V. Estab. Pat. (90270)	16.67 1.40 (11.00)	34.18 2.12 (14.00)
12. Extended O.V. Estab. Pat. (90070)	17.85 1.50 (12.00)	26.71 1.65 (11.00)
13. Brief Home V. Estab. Pat. (90140)	18.08 1.52 (13.00)	20.92 1.30 (7.00)
14. Limited Home V. Estab. Pat. (90150)	20.04 1.69 (14.00)	24.54 1.52 (9.00)

Table 4.3 (Continued)

	Time-Based[a]	Charge-Based[b]
15. Arthrocentesis (20610)	24.12 2.03 (15.00)	25.62 1.59 (10.00)
16. Comprehensive O.V. New Pat. (90020)	25.68 2.16 (16.00)	50.71 3.14 (19.00)
17. Comprehensive O.V. Estab. Pat. (90080)	26.98 2.27 (17.00)	43.25 2.68 (17.00)
18. Comprehensive H.V. New Pat. (90220)	37.30 3.14 (18.00)	60.18 3.73 (20.00)
19. Herniorrhaphy (49505)	41.82 3.52 (19.00)	439.38 27.2 (21.00)
20. Thoracentesis (32000)	44.11 3.71 (20.00)	49.75 3.08 (18.00)
21. Hysterectomy (58265)	56.13 4.73 (21.00)	861.21 53.3 (24.00)
22. Cholecystectomy (47600)	59.62 5.02 (22.00)	705.65 43.7 (23.00)
23. Proctosigmoidoscopy (45300)	62.00[c] 5.22 (23.00)	42.34 2.62 (16.00)
24. Colon Resection (44140)	98.25 8.27 (24.00)	972.29 60.20 (25.00)
25. Heart Catherization (93527)	120.92 10.20 (25.00)	532.28 33.00 (22.00)
All Procedures		
Mean	2.65	10.16
Standard Deviation	2.35	17.93

Notes: a. Mean time per procedure in minutes.

b. Mean HCFA (unindexed) prevailing charge in dollars.

c. Based on data for general surgeons only.

shows that the surgical procedures have charge-based relative values about 6.5 times greater than their corresponding time-based relative values. Charge-based relative values are about 50 percent greater for hospital visits, 20 percent greater for office visits, and 10 percent lower for the remaining procedures in the 2 scales.

Table 4.4 reports Pearson and Spearman correlation coefficients among the scales. Separate correlations were computed for the visit and nonvisit procedures. The Pearson correlation among all visits (office, hospital, and home) in the 2 scales is quite high, 0.88. For the nonvisit procedures, in contrast, it is 0.65. But this understates the differences in the scales for the nonvisit procedures because the time-based scale has a mean of 4.09 compared to 22.78 for the charge-based scale.

The fact that the ratios of charge-based to time-based relative values are not the same for each procedure group suggests that the two scales are substantively different, i.e., one is not a simple scalar multiple of the other. Put another way, time spent performing different types of procedures is valued differently by a charge-based system. Factors noted earlier--skill, complexity, training, and outcome--are presumably part of the story. Differences in insurance coverage for different types of procedures may also be important. For example, a hospital visit is more likely to be covered by insurance because a hospitalized patient is more likely to have satisfied the annual deductible than the average patient seen in the office.

Another factor influencing charge-based relative values is differences in interphysician competition by procedure type. Again, looking at relative values for office and hospital visits, competition among physicians and patients' price sensitivity are probably greater for office visits than for hospital visits. Once the patient is hospitalized, the physician is much

Table 4.4

Correlation Between Time-Based and Charge-Based RV Scales[a]

	Visits (N=15)	Other Procedures (N=10)
Pearson Correlation Coefficient	0.88	0.65
Spearman Correlation Coefficient	0.83	0.75
Scale Mean		
Time-Based	1.40	4.09[b]
Charge-Based	1.74	22.78
Scale Standard Deviation		
Time-Based	0.69	2.77[b]
Charge-Based	0.91	23.67

Notes: a. Scales are constructed from mean times and mean HCFA prevailing charges per procedure.

b. 13 procedures; includes I & D, lumbar puncture, and mastectomy, which are not included in HCFA's prevailing charge data base.

Table 4.5

Characteristics of Office and Hospital Visits

Procedure (CPT-4 Code)	Characteristics									
	No. of Encounters	Severity[a]				Urgency[c]				
		None	Minor[b]	Moderate[b]	Severe[b]	None	Deferrable	Same Day	Sooner	Emergency
1. Brief Office Visit, Established Patient (90040)	26,843	15.9%	57.6%	24.0%	2.2%	19.3%	29.8%	48.2%	1.5%	0.8%
2. Brief Hospital Visit, New Patient (90200)	3,949	18.5	19.6	46.6	14.9	11.3	8.3	71.0	2.9	5.5
3. Limited Office Visit, Established Patient (90050)	29,736	16.6	36.6	42.4	4.3	19.0	23.6	53.4	2.3	1.2
4. Limited Hospital Visit, Established Patient (90250)	7,905	6.3	13.0	56.1	24.6	5.1	7.2	82.5	1.8	2.9
5. Extended Office Visit, Established Patient (90070)	3,595	7.1	12.7	61.1	18.8	12.3	26.7	53.0	4.6	2.6
6. Extended Hospital Visit, Established Patient (90270)	4,571	1.1	3.8	52.2	42.8	2.5	6.8	72.1	7.1	11.0
7. Comprehensive Office Visit, New Patient (90020)	1,485	18.65	18.0	48.1	14.7	22.8	23.0	33.3	10.4	10.0
8. Comprehensive Office Visit, Established Patient (90080)	1,246	12.84	16.3	44.5	26.2	22.4	26.3	41.7	6.0	2.3
9. Comprehensive Hospital Visit, New Patient (90220)	1,088	3.95	4.1	30.3	61.3	3.0	5.1	46.5	9.7	35.5
All Visits	80,418	14.0	36.4	39.2	10.2	16.0	22.3	55.8	2.7	2.7

Notes: a. Percentage distribution of encounters in each procedure by severity. See Figure 4.4, USC-Mendenhall Coding Key.

b. Combines acute and chronic conditions.

c. Percentage distribution of encounters in each procedure by urgency. See Figure 4.4, USC-Mendenhall Coding Key.

d. Percentage distribution of primary specialties of physicians providing encounters.

Table 4.5 (Continued)

Procedure (CPT-4 Code)	Characteristics					
	Primary Specialty of Physician[a]					Pct. of Board-Certified Physicians
	Fam. Prac.	Gen. Prac.	Gen. Surg.	Int. Med.	Ped.	
1. Brief Office Visit, Established Patient (90040)	40.0%	12.1%	5.7%	15.6%	23.9%	9.1%
2. Brief Hospital Visit, New Patient (90200)	18.6	4.6	12.8	33.9	24.9	19.2
3. Limited Office Visit, Established Patient (90050)	31.4	7.4	3.5	19.2	35.7	12.1
4. Limited Hospital Visit, Established Patient (90250)	21.4	7.3	13.0	41.3	12.0	28.4
5. Extended Office Visit, Established Patient (90070)	31.2	12.3	6.2	31.4	12.8	20.5
6. Extended Hospital Visit, Established Patient (90270)	22.6	10.7	15.1	37.5	8.5	20.8
7. Comprehensive Office Visit, New Patient (90020)	27.5	7.7	8.2	29.4	20.4	16.3
8. Comprehensive Office Visit, Established Patient (90080)	24.1	7.7	2.8	47.6	12.9	33.5
9. Comprehensive Hospital Visit, New Patient (90220)	15.4	3.4	6.3	50.9	14.3	26.7
All Visits	31.8	9.2	6.5	23.5	25.4	14.5

closer to being a monopoly provider. As a result of these market forces, charges for time spent providing office visits would be expected to be lower than charges for time spent providing hospital visits, even if the physician's other costs may be higher in the office than in the hospital.

Using other information from the USC-Mendenhall survey, we can examine some of the differences in the patients and physicians involved in each of the CPT-4 visit codes. As shown in Table 4.5, hospital visits in each duration category (brief, limited, extended, and comprehensive) have higher severity and urgency ratings than office visits of the same nominal duration. In addition, the mix of physicians providing hospital and office visits is not the same. General surgeons and internists, and board-certified physicians provide higher proportions of the hospital visits than of the office visits. These data suggest that the higher relative values of hospital visits in a charge-based system may be partially due to the greater difficulty of the cases treated and the greater training of the physicians conducting the visits. As noted earlier, these are factors that time alone does not capture very well.

D. SUMMARY AND DISCUSSION

Construction of time-based RV scales from data on physicians' time per procedure is straight-forward mechanically. However, such scales have serious theoretical and implementation problems. The biggest theoretical problem is that physicians' time is not combined in fixed proportions with all of the other resources that go into producing physicians' services. As such, relative physicians' time would be a poor indicator of procedures' relative costs. In principle, time estimates could be adjusted to correct for variations in procedure complexity, the contributions of other inputs, and physicians' investments in their own training. In practice, such adjustments require making several arbitrary assumptions as well as substantial data on practice-related factors other than physicians' time.

The biggest problem in actually implementing a time-based RV scale is the lack of a sufficiently large data base which reports time estimates for procedures identified by their CPT-4 codes (or an equivalent coding scheme). Mounting and managing a survey of adequate size would be a substantial expense. Furthermore, the reliability of the times reported might be distorted if it were known that the purpose of the survey were to construct an RV scale.

Using information available from the USC-Mendenhall study of physicians' practices, we identified 28 procedures that appeared to be nearly equivalent to CPT-4 procedure codes. Examination of RV scales constructed for these procedures suggested that relative values were fairly robust with respect to the choice of a distribution point (mean, median, etc.) to represent a procedure's absolute value.

Subsequent comparisons between a time-based and a charge-based RV scale (constructed from mean times and mean HCFA prevailing charges) suggested that there are substantial differences between the two in the valuation of surgical procedures--their charge-based relative values were about 6.5 times greater than their time-based relative values. This should not be surprising, since the time data were not adjusted to reflect complexity, outcome value, or other costs. However, there is no clearcut method for adjusting the time data. Nor is there any mechanical or statistical method for assessing whether one set of relative values is too high or the other set too low. Differences between time-based and charge-based relative values for visits were smaller. Charge-based values were about twenty percent higher for office visits and about fifty percent higher for hospital visits.

CHAPTER 5

THE MICRO-COSTING APPROACH

by

Judith Wagner

A. DEFINITION OF MICRO-COSTING

Micro-costing is an approach to measuring the resource costs of individual products or services when these are produced in common with other products or services. In a multi-product manufacturing firm, for example, the equipment, materials and personnel may be shared in varying proportions in the production of several products. It is important both for pricing and production decisions that managers have accurate estimates of the cost of producing each product. Micro-costing employs the techniques of time study, work sampling and cost accounting to allocate costs to specific outputs of the firm. (Shuman, Wolfe and Perlman, 1973) It is a "bottoms-up" approach in the sense that certain inputs, such as labor and equipment, are directly observed in the production of each kind of product. The time required for each class of employee to engage in the production of a unit of each product is estimated via direct observation of the production process. Similarly, the amount of time on major equipment required for each product might be determined by direct observation. Costs which are not directly related to the production of specific products, such as idle time, rent, overhead, etc., are allocated by principles that most closely approximate their relative contribution to the different products. For example, the cost of idle time of a specific worker class might be allocated among the different products according to the proportion of the workday spent on each product by that class of workers. (Shuman, Wolfe and Perlman, 1971)

The essence of micro-costing is direct observation of the production process through time studies and work sampling. (Barnes, 1958) These studies involve direct observation and analysis of data using straightforward statistical sampling techniques. A time study involves continuous monitoring of the personnel and/or equipment associated with the production of a sample of items produced by the firm. In work sampling, a sample of employees is observed at randomly selected times in the workday. The former type of study gives estimates of the mean time associated with the production of each product by each class of worker, while the latter estimates the proportion of time spent by each class of worker in all work-related activities, those directly productive and nonproductive.

Together, these two kinds of observational studies give data on which the allocation of costs can proceed. Micro-costing goes beyond time and work sampling studies, assigning dollar costs to each product arising from the direct and indirect inputs into the production process. The result of a micro-costing study is a set of average unit costs for each final product of the firm.

Because micro-costing is based on primary data collection techniques and, in particular, on direct observation of the production process, it is a relatively expensive endeavor. (The cost of such studies in health care applications will be discussed later in this chapter.) Thus, it can only be justified in situations where there are several shared inputs used in very different proportions in the production of multiple products.

B. MICRO-COSTING AND RVSs IN HEALTH CARE

Because it is most useful when non-physician personnel and equipment are important inputs into the production process, micro-costing is best applied in equipment-intensive areas of health care, such as clinical laboratory, radiology and specialized medical procedures. In the remainder of this paper, then, we consider the use of micro-costing in the development of relative values for clinical laboratory, radiology, and other special equipment-intensive services offered by physicians or other providers.

The transformation of a given set of unit costs into a relative value scale (RVS) is trivial. After selecting one procedure as the numeraire, relative values are computed by taking the ratio of other procedures' unit costs to the numeraire's unit cost.

To assess the validity and relevance of such a scale, it is important to understand the asumptions underlying a micro-costing-based RVS. They are as follows:

- The observed level of utilization of capacity (both equipment and personnel) is optimal.

  As a cost accounting technique, micro-costing accepts the observed levels of idle or other unproductive time as unavoidable costs of production. Yet, these measured quantities will depend upon the institutional environment in which measurement takes place. If a department is operating below full capacity, its measured unit costs will be high, and vice versa. This is a problem for RVS construction only if excess capacity falls disproportionately on services so that relative as well as absolute costs are affected by the level of capacity utilization. For example, if a particular piece of equipment used in a subset of procedures is not used to capacity, the cost of those procedures will be overestimated relative to that of others.

When some services must be provided on demand, there are economies of scale in the use of resources. A high-volume organization needs less proportional idle time to maintain a given level of service than does low volume organization. Thus, measurement of costs in a high-volume environment would underestimate the costs of idle time for a low-volume facility. Rural facilities, for example, are likely to require more excess capacity to maintain ready availability of certain services than would urban facilities.

- The organization and technology of the observed setting(s) is optimal.

The setting(s) selected for study dictate the level of technology and the organization to be observed. Rapidly changing technologies can render a given study at least partially obsolete. For example, in the area of clinical laboratories, automated equipment has drastically reduced the amount of labor required to process certain chemistry tests. (Penner, 1982) Though the equipment costs are high, there is clear evidence that automation has reduced the cost of testing at high volumes. (Fineberg, 1978) Relative values based on preexisting technologies would not reflect this new efficiency.

The organizational environment is equally important to the resulting relative values. The personnel mix in a laboratory, for example, will determine the relative costs of testing. One study of reagent preparation in clinical laboratories showed that the relative costs of commercially prepared kits versus in-house preparation depend largely on the testing volume of the laboratory. (Richardson, 1977) Even the physical layout of the facility can affect the observed cost of operation.

- <u>Proficiency in the observed setting is optimal.</u>

  This issue is particularly important for new procedures, which may be costly to perform when they are unfamiliar but which may become faster and cheaper as those involved gain proficiency in their production. Measurement of costs too early in a procedure's use can lead to an overstatement of costs as they evolve.

- <u>The observed quality of service is optimal.</u>

  Some medical services can be produced more quickly and cheaply only at the expense of quality. A radiographic examination, for example, can vary widely in quality and may in part be a function of the time and care given both by the technician and the reader. Also, if the relative value scale does not separately account for examinations with different numbers of views, then the examination cost would reflect the mix of views actually used in the setting under study. This mix may not be optimal in other settings.

The strictness of these assumptions helps explain why RVSs based on micro-costing studies would be sensitive to the selected sample of organizations on which they are based. The replicability of studies across settings and over time would be low because of the importance of minor changes in the assumptions listed above. It also explains why multiple studies in a wide range of settings would be necessary for sufficient information on which to base scale construction. This brings us to consideration of the cost and practicality of RVS construction using micro-costing techniques.

C. THE COST AND PRACTICALITY OF RVS BASED ON MICRO-COSTING

Several micro-costing studies have been conducted in hospitals, generally within equipment-intensive departments such as clinical laboratory, radiology and surgery. These departments also rely heavily on non-physician personnel in the production of their services. It is estimated that a complete micro-costing study of a single hospital radiology department would cost about $500,000 today. (Personal communication with Harvey Wolfe) A mid-1970's micro-costing study of thirteen common minor surgery procedures in six facilities in a city in the southwest cost over $200,000. (Personal communication with Franz Jaggar, The Orkand Corporation) A British researcher estimated that a single-facility micro-costing study of sixty clinical laboratory procedures in 1980 cost almost $200,000. (Personal communication with A. Stilwell) Thus, the cost of micro-costing is high, even when the locus of observation is limited. More general, periodically updated, analyses across a large number of settings would be prohibitively expensive.

Relying as it does on direct observation of the production process, micro-costing is an intrusive approach to RVS construction. Issues of confidentiality and patients' privacy arise and complicate the design of studies.

There is also a problem of gaining access to varied settings in which such studies might be conducted. Those who are willing to cooperate may represent a biased sample of providers. Also, the very fact of being observed may alter the production process in ways that cannot be anticipated.

In the case of clinical laboratory procedures, there already exists a data base consisting of non-physician personnel time requirements for each procedure with each avilable type of equipment. This data base was developed by the College of American Pathologists* as a method for improving labor utilization in laboratories. (College of American Pathologists, 1982) It is derived from time studies submitted voluntarily by laboratories across the United States and represents the average experience in these settings. The time requirements are updated and published annually in the CAP's Laboratory Workload Recording Method (WRM) manual. Since the data represent the contribution of many clinical laboratories and are available to all users, it is tempting to consider the WRM for use as a relative value scale. It is important to recognize, however, that the WRM units are not relative costs; they merely represent the relative input of non-physician laboratory technical time into the production of specific laboratory procedures. The costs of equipment, commerical preparations and reagents, which may vary markedly across procedures, are not included, nor is the cost of the pathologist's time, which is generally high in certain procedures such as cytology and negligible in chemistry procedures. (Glenn, 1982) Thus, the WRM is a non-physician time-based, not a cost-based, relative value system.

Nevertheless, it is worth considering whether this data base could be used in the construction of a micro-costing based RVS, thereby avoiding the

---

*The system was originally developed by the Canadian Association of Pathologists.

costs of direct observations. Unfortunately, micro-costing would still require work sampling in order to measure and allocate pathologists' work times and technicans' idle times. Thus, the availability of the WRM data does not obviate the need for expensive and intrusive observational studies.

D. EXAMPLES OF MICRO-COSTING STUDIES

In this section we present the results of some micro-costing studies and compare them with relative value scales published in the 1974 California Relative Value Studies. Although there may be many applications of micro-costing in the field, results have been published only sporadically. All of the applications described here have been performed in hospitals or similar institutions. All are specific to a particular department. Therefore, the examples are organized by department.

Radiology

In 1972, Shuman et al. reported on the results of a micro-costing study of radiology procedures in two large teaching hospitals, one specializing in adult care and one in pediatrics. (Shuman et al., 1972) The average cost of each of 63 procedures was estimated from a comprehensive work sampling study in each hospital.

The radiologist times and total unit costs of 63 procedures reported by Shuman et al. were converted to relative value scales for use here* as shown in Table 5.1. These relative values are also compared to the total procedure value and the professional component given to each of the 63 procedures by the

*The sialogram was selected as the numeraire in all scales because it is unambiguously defined. The chest x-ray, the most common procedure, can vary widely in number of views and is therefore imprecisely defined. All other procedure values in a scale were divided by the value of this procedure in that scale.

Table 5.1

Relative Values of Radiology Procedures

Procedure Name (CPT-4)	1974 CRVS[1] Physician Component	1974 CRVS[1] Total Unit Value	Shuman & Wolfe[2] Radiologist Time	Shuman & Wolfe[2] Total Unit Exam Cost	Brouhard[3] Technical Staff Time
GROUP I					
Abdomen (74000, 74020)	.84	.76	.28	.31	—
Chest (71020, 71030)	.81	.86	.28	.31	.13
Extremeties	—	—	.28	.36	—
Kidney, Urether and Bladder (74400)	1.87	1.84	.28	.31	—
Hips (73500, 73510)	.68	.74	.28	.31	—
Pelvis (72170, 72190)	.75	.79	.28	.26	—
Ribs (71100, 71110)	.89	.96	.28	.36	—
Shoulder (73030)	.71	.75	.28	.31	—
Spine (T) (72070)	.77	.86	.28	.31	—
Spine (LS) (72110)	1.35	1.45	.28	.52	—
Spine (C) (72040)	1.07	1.13	.28	.47	—
Bone Survey (76060)	1.81	1.82	.46	.99	—
Facial Bones (70140)	.77	.86	.28	.52	.40
Madibles (70100, 70110)	.91	.96	.28	.36	.40
Nasal Bone (70160)	.71	.76	.28	.43	.40
Orbital Cavity (70400)	BR[a]	BR[a]	.28	.41	—
Sinuses (76080)	1.19	1.06	.28	.47	—
Skull (70250, 70260)	.97	1.07	.46	.63	.26
Laminogram (76100)	1.74	1.61	.28	.62	—
Sacrum (72220)	.74	.80	.28	.31	—
Coccyx (72220)	.74	.80	.28	.31	—
Sacro Iliac Joint (72202)	.90	1.00	.28	.31	—
Sternum (71120)	.71	.75	.28	.36	.33
Sternovascular Joint (71130)	.71	.75	.28	.31	—
Humerus (73060)	.59	.61	.28	.31	—
Elbow (73070, 73080)	.63	.65	.28	.47	—
Forearm (73090)	.59	.59	.28	.47	—
Wrist (73100, 73110)	.60	.62	.28	.41	—
Hand (73120, 73130)	.57	.59	.28	.31	—
Finger (73140)	.45	.45	.28	.36	—
Femur (73550)	.71	.75	.28	.31	—
Knee (73560, 73570)	.65	.66	.28	.47	—
Tibia (73590)	.59	.61	.28	.31	—
Ankle (73600, 73610)	.60	.62	.28	.36	—
Foot (73620)	.49	.58	.28	.36	—
Toe (73660)	.40	.45	.28	.31	—
Os Calcis (73650)	.55	.53	.28	.31	—
Optic Foramina (70190)	.71	.75	.28	.31	.40
Zygomatic Arch	—	—	.28	.37	—
Tempomandibular Joint (70330)	1.03	1.10	.28	.41	.40
GROUP II					
Barium Enema (74270)	1.74	1.47	1.00	1.31	—
Bronchogram (71040, 71060)	2.38	2.17	1.00	1.42	—
Chest Floro (76000)	.71	.51	1.08	.95	.40
Cholocystogram (74290)	1.13	1.18	1.00	1.05	—
T-Tube Cholengiogram (74305)	1.55	1.41	.82	1.11	—
IVC (74300)	1.29	1.49	1.00	1.16	—
Esophagus (74220)	1.19	1.06	1.08	1.08	—
Myelogram (72250, 72270)	3.23	2.89	.82	1.53	—
IVP (74400)	1.87	1.84	.82	1.16	—
Small Bowel (74250)	1.81	1.67	1.00	1.21	.40
Sinogram (70210, 70220)	.80	.93	1.00	1.05	—
Sielogram (70390)	1.00	1.00	1.00	1.00	1.00
Venogram	—	—	1.00	1.16	—
GI Series (74240)	1.74	1.70	1.00	1.34	—
Retrograde Pyelogram (74420)	1.35	1.47	1.00	1.05	—
Ventriculogram	—	—	1.00	1.16	—
Cystogram (74430)	1.00	1.08	1.00	1.10	—
GROUP III					
Aortogram	—	—	15.09	16.79	3.33
Arthogram	1.29	1.25	14.79	16.79	—
Carotid Angiogram (75680)	4.19	—	15.09	16.79	—
PEG (70002)	1.90	4.90	8.30	8.32	—
Lymphangiogram (75802, 4,6,8)	3.87	4.51	11.74	12.13	—

Notes: a. BR--By report.

Sources: 1. California Relative Value Studies, 1974 Revision. Prepared by the Committee on Relative Value Studies, California Medical Association, 1975.

2. Shuman, L.J., Wolfe, H., Ghaly, R. and Palaniappan, P. Manual for Implementing Radiology Micro-Costing. Department of Industrial Engineering, Systems Management Engineering and Operations Research, University of Pittsburgh, August 1972.

3. Brouhard, Chris E. "System Monitors Productivity in Diagnostic Imaging," Hospitals 55(4): pp. 167-175, July 16, 1981.

1974 California Relative Value Study (CRVS). Finally, in a separate time study Brouhard reported on the average technician time for 12 procedures (Brouhard, 1981) in a single institution. Table 5.1 includes these relative values as well.

The correlation among the scales was studied in order to assess the degree to which they are reasonable substitutes for one another. The zero-order product moment correlation between the physician's component of the CRVS and the radiologist's time was 0.62 (p .01), indicating that, while there is a strong positive relationship between the values assigned in the scales, they are not exact substitutes for one another. The correlation between the CRVS total unit value and the unit cost was 0.50 (p .01), again indicating that there is substantial difference in the two scales.

As might be expected, the scale based on technician's time alone bore little relation to either CRVS scale (correlation coefficients were not significantly different from zero) and was only mildly correlated with the total unit cost scale (r = 0.46, p = .12).

It is interesting to note that the correlation between total unit cost and radiologist time, both measured by Shuman et al., is extremely high (r = 0.998). This is due both to the importance of the radiologist as a high priced input and the techniques used to allocate indirect costs. Most of the indirect costs were allocated according to examination time. This may imply that a relative value scale based on radiologist time studies is a reasonable substitute for an RVS based on a full micro-costing analysis, if the assumption used to allocate indirect costs is appropriate.

It is unfortunate that we do not have two or more separate micro-costing studies in order to determine how stable the scales based on such studies are likely to be. Since the CRVS does not purport to represent relative costs, the variation between the two scales is not surprising. Yet the professional component of the CRVS should reflect to some extent relative time requirements. Further scrutiny reveals that the micro-costing study produced a much wider range of radiologist time values (.28 - 15.09) than did the CRVS professional component (0.40 - 4.19). Shuman, Wolfe and Perlman have suggested that the CRVS professional component scale may be weighted in favor of routine procedures at the expense of the infrequent but time-consuming special procedures (Shuman, Wolfe and Perlman, 1971). Of course, such judgments are premature, since the micro-costing study was based on only two teaching hospitals. Micro-costing-based scales in private radiologists' practices might have different results.

Clinical Laboratory

Micro-costing has been applied to several areas of the clinical laboratory, including microbiology (Shuman, Wolfe and Perlman, 1971; Wild, 1974) and clinical chemistry (Stilwell, 1981). Here we compare the results of the clinical chemistry study with CRVS values and with the College of American Pathologists' WRM values.

Stilwell examined the costs of sixty common chemistry tests in a British hospital and reported on the costs of five such examinations. (Stilwell, 1981) The direct and indirect costs of each procedure were estimated via work sampling and other allocation rules. Table 5.2 shows the relative values of each test. The range of testing costs is much narrower than the range given in the 1974 CRVS. Comparing Stilwell's results with the 1982 CAP workload values demonstrates the variations between cost-based and time-based scales. Chemistry profiles (which given parallel findings on several chemicals in a single blood serum specimen) are highly automated procedures involving small amounts of technologist time but high equipment costs. The 1982 CAP unit

Table 5.2

Relative Values: Clinical Laboratory

Procedure Name (CPT-4)	1974 CRVS[1] Unit Value	Stilwell[2] Direct Costs	Stilwell[2] Total Costs	1981 CAP[3] Unit Value
Overall Costs (80010) Profile (10 elements)	2.70	1.05	1.39	SMA[a] .60 SMAC[b] .25
Glucose (82947,82948)	.62	.72	.85	.80
T-3 Uptake (84250)	.95	.72	.80	--
Acid Phosphatase (84060)	1.00	1.00	1.00	1.00
Electrophoretic Strip	--	.78	.96	1.20
Range (Max. - Min.)	2.1	.33	.59	.40 - .75

Notes: a. Technicon's SMA Automated blood chemistry analyzer.

b. Technicon's SMAC blood chemistry analyzer.

Sources: 1. California Relative Value Studies, 1974 Revision. Prepared by the Committee on Relative Value Studies, California Medical Association, 1975.

2. Stilwell, A. "Costs of a Clinical Chemistry Laboratory." Journal of Clinical Pathology 34: 589-594, June 1981.

3. Laboratory Workload Recording Method, 1982 Edition. Prepared by Workload Recording Committee, College of American Pathologists, 1982.

value for a profile ranges from 0.25 to 0.60, depending on the equipment, compared to a value of 1.39 in Stilwell's study. Interestingly, the CRVS value for a profile test is almost twice as high as the equivalent cost-based value.

Glenn has proposed the use of CAP WRM values in the construction of cost-based fee-setting for new procedures undertaken by a laboratory. (Glenn, 1982) The CAP units would be used in lieu of time studies. In this system, indirect costs would be allocated to the new procedures in the same proportion of direct costs as they occur in the laboratory as a whole. Though the approach is approximate, Glenn contends that this modified micro-costing approach would give a better approximation of the true costs of performing a new procedure than would any other feasible approach to estimating costs. Thus, the availability of the CAP WRM values, while insufficient for a comprehensive micro-costing study, does make cost analysis of individual new procedures relatively attractive.

Obstetrics/Gynecology (OB/GYN)

The use of micro-costing as a basis for RVS development of hospital-based OB/GYN is ill-advised, since physicians are typically paid separately from hospitals. An RVS for physician fees should not include the costs borne by the hospital. But a micro-costing study of six common OB/GYN procedures performed by MacDonald and Reuter also reported on the results of a time study for physician personnel in the performance of six hospital-based OB/GYN procedures in a large teaching hospital. Table 5.3 summarizes the results of that study and also gives the equivalent 1974 CRVS values. There is a wide disparity between physician times as reported by MacDonald and Reuter and the CRVS unit values. This disparity, particularly in the rank-order of procedure values, is greater than is the disparity between the 1974 CRVS values and the

Table 5.3

Relative Values, OB/GYN

Procedure (CPT-4)	1974 CRVS[1]		MacDonald and Reuter[2]			
	Unit Value	Anesthesiologist Time	Physician Time	Anesthesiologist Time	Nurse/Technician Time	Total Unit Costs (Average Capacity Utilization)
Total Abdominal Hysterectomy (58180, 58200)	3.79	2.00	1.36	1.00	1.34	6.88
Saline Injection	--	1.00	1.44	--	1.54	2.13
Normal Delivery (59400)	2.26	1.09	1.37	1.15	1.30	2.86
Caesarean Section (59520)	2.93	2.00	1.03	1.00	1.20	4.89
Laparoscopy (58982, 58983)	1.56	1.85	1.03	1.15	1.10	1.83
Dilation & Currettage (58120)	1.00	1.00	1.00	1.00	1.0	1.00

Sources: 1. California Relative Value Studies, 1974 Revision. Prepared by the Committee on Relative Value Studies, California Medical Association, 1975.

2. MacDonald, L.K. and Reuter, L.F., "A Patient Specific Approach to Hospital Cost Accounting." Health Services Research, Summer 1973: 102-121.

relative total costs. If the total costs of a procedure are viewed as an index of the risk and severity of the conditions occasioning it, the CRVS values may reflect the differential skill and risk required of each, whereas relative physicians' times do not.

E. SUMMARY

An examination of three RVSs based on micro-costing techniques reflects substantial differences from the California Relative Value Scale. In radiology, neither unit costs nor physician times correlated highly with the CRVS values, but physician times as measured in the micro-costing study appeared to be a good surrogate for unit costs.

In the laboratory, unit costs, especially for automated procedures, were very different from those of the CRVS and also deviated substantially from the CAP workload units.

In OB/GYN, the rank-order of unit costs agreed with the rank-order of CRVS values, but the rank-order of physician times as computed in the cost study did not. Since physicians do not bear the hospital costs, the relative values in this area should not reflect unit costs, but the agreement between CRVS and unit costs may reflect an underlying assessment of severity, skill and risk associated with the medical procedures.

Evaluation of Micro-Costing as an RVS Construction Method

In this section, we summarize the findings regarding the performance of micro-costing with respect to the eight criteria set forth earlier in this report.

Data Requirements

The need for primary data collection at the source of medical care delivery renders the micro-costing approach an intrusive and expensive approach to scale construction. For office-based procedures, the cost would be prohibitive because separate studies would need to be performed in a large number of physicians' offices. Even then, the reliability of the resulting estimates is likely to be low.

Technical Expertise Required

Micro-costing studies require the involvement of a statistician or industrial engineer for the design and evaluation of the observational studies. Clerical staff could be trained for data collectors. Accounting or business expertise would also be useful for development of rules for allocating indirect costs.

Ease of Development of Specialty- and Location-Specific Scales

Since a micro-costing RVS is likely to be extremely sensitive to relatively small variations in practice patterns across specialties and locations, this approach would require replication in each specialty and location deemed appropriate for a separate scale. In light of the costs of direct observation, this method is of restricted usefulness for development of a family of specific scales.

Potential for Future Updating of Scales

Once a scale is constructed, it may become invalid due to changes in technology or in input prices. Changes in the latter could be accommodated rather easily by recomputation of unit costs using new input prices, provided that the change in input prices does not change the mix of inputs. Changes in technology, however, could be incorporated into a scale only by repeating direct observational studies. In areas of rapid technological change, such as laboratory and radiology, periodic reconstruction of the scales would be necessary at high cost.

Pricing Problem Medical Procedures

Micro-costing is a feasible approach to establishing an initial relative value for new medical procedures. Although the cost is high because of the need for direct observation of practices, it would be possible to select a few settings in which the volume of a particular procedure is high enough to insure adequate technical proficiency and high capacity utilization. In the clinical laboratory, micro-costing for new procedures could build on the CAP workload units reported in the field, thus obviating the need for time or work sampling studies.

Suitability Across Procedures

In principle, there is no barrier to the application of micro-costing across broad classes of procedures. In practice, however, the reliability of such inter-class comparisons is likely to be low, even lower than that of intra-class scales, because the setting of care varies so widely among procedure categories.

External Costs

Micro-costing cannot account for costs outside of the setting in which cost-measurement takes place. It is a resource-cost concept and is not intended to measure costs to society.

CHAPTER 6

GROUP DECISION-MAKING METHODS

by

Robert Berenson

As pointed out elsewhere, available data on which to base a relative value schedule are for the most part confined to charge-based information. There are no comparable bodies of reliable data for cost-related information, such as time per procedure, complexity, severity, resource costs, etc. Social preferences by definition cannot be based on a simple data set or a particular analytic technique.

In the absence of an accepted data base which reflects time, costs, preference or other noncharge bases for measuring the value of physicians' services, RVSs can be relatively easily established using charge data. As far as we have been able to determine, all previous large-scale attempts in the U.S. to set up RVSs, whether by insurance companies or by medical societies, have been based on charge-related information, usually deriving from the updates of the California RVS.

Nevertheless, it is conceptually possible to develop RVSs based on factors other than charges, even in the absence of a precise data base or use of a particular analytic technique. In health care at all levels, unstructured conferences and meetings are commonly employed to establish standards of care and criteria for assessments of various kinds. Medical societies in the past have convened panels of physicians to examine established charge-based RVSs in order to assign values to new procedures or make marginal adjustments based on implicit "cost" considerations. This unstructured approach could be extended to RVS development. As this chapter will demonstrate, however, a process as complex and important as national RVS setting would require a structured approach which recognizes the importance of group representation, the deliberative process, and data needs.

A. METHODS OF STRUCTURED DECISION-MAKING

There are structured methods available to utilize informed opinion as a substitute for precise information in order to reach consensus on relative values of procedures. These group decision procedures would permit more rigorous consideration of time and other cost factors. They would not be as satisfactory for consideration of social preferences, as will be made clear below. The essence of the consensus development approach would be the use of a panel of experts whose pooled judgments represent a "best guess" about how services do, in fact, relate to one another.

Structured methods for achieving group decisions utilizing expert opinion include "consensus development" as used by the National Institutes of Health, the Delphi technique, and the Nominal Group Technique (OTA, Strategies Medical Technology Assessment, September 1982).

Consensus Development

"Consensus development" has been adopted by NIH to seek consensus on the safety, efficacy, and appropriate conditions for use of various medical technologies. In this process, a panel of experts drawn from various disciplines is selected by NIH to hear presentations by researchers addressing a prespecified set of questions about the technology. After hearing from the researchers, the panelists and audience members discuss the findings. The panel, usually of 10-20 members, then withdraw to deliberate for the purpose of achieving a consensus statement. In rare instances, minority reports are developed to indicate disagreement with the majority recommendations. After soliciting the views of the audience on a draft consensus statement, the panel revises the statement for dissemination. The statement is supposed to identify what is known and not known about the technology (Perry and Kalberer, 1980). This process, familiar to physicians in their work on hospital utilization review and audit committees, is a face-to-face process of give and take with loosely structured rules for interaction.

Delphi

The oldest structured model and best known approach for involving groups in decision-making is the "Delphi Method." Although much less commonly used than unstructured face-to-face committee approaches, Delphi has been used to help solve certain health policy problems. In contrast to simple questionnaires administered to a group of respondents, Delphi involves three to four successive rounds of anonymous questionnaires with feedback of information to expert respondents between rounds. In the first round, for example, individuals might be asked to assign relative values to a sample of procedures with only very general instructions. The median rating could then be returned to the respondents along with additional information, e.g., charge-based relative values, with a request to perform a second rating. More elaborate Delphi questionnaires could ask respondents to list the considerations underlying their rating, and these qualitative judgments could be fed back to all respondents in successive rounds.

Alternatively, Delphi groups of experts could be provided questionnaires with very precise instructions for answering specific questions, e.g., how much time on average do physicians spend to perform a list of procedures. Again, feedback of the group and individual responses would be provided to respondents over three or four rounds in order to narrow the range of responses to finally achieve "consensus."

Experiments with the Delphi technique in forecasting tasks have shown that convergence to a consensus is common over three to four rounds of questionnaires, where consensus is defined as an acceptably low level of variation in estimates around the mean or median (Dalkey, 1968). Consensus is not assured under Delphi for essentially judgment tasks where there is no potentially verifiable solution.

A unique feature of the Delphi technique is that persons selected to participate in the process generally have no direct contact with one another. Instead, participants are provided with a summary of the questionnaire responses, usually by mail. Personality or status variables thus have little chance to influence participants' opinions, as they might in face-to-face meetings. By using anonymous feedback, each participant has an equal chance of influencing other participants (Bunning, 1979).

Nominal Group

"Nominal group technique" (Delbecq et al., 1975) fits between "consensus development" and "Delphi." Like "consensus development," participants do sit around a common table and are expected to achieve consensus within a strictly limited time frame. The panel is asked to write their views on the issues posed by the leader of the meeting. But each person's view is recorded on a separate card and talking is prohibited. The cards are collected and their contents are listed for all to see. However, like Delphi, the authors of the individual cards are not identified; and, therefore, domination of the group thinking is theoretically minimized.

Whether Delphi performs better than face-to-face groups is subject to conflicting evidence. Campbell (1968) found Delphi groups to be more accurate at short-term economic forecasts than face-to-face encounters, but Farquhar (1970) found significantly better results in face-to-face groups required to solve a complex estimation task. However, neither of these studies compared the impact of group size on the accuracy or quality of the outcome of either kind of group. Because Delphi allows for participation of more experts than do face-to-face groups, which have a practical limit of about 15 (Filley, 1970), it may in reality be a more accurate technique. The anonymity of the Delphi process may be another advantage compared to face-to-face meetings, since it lessens interspecialty conflicts over medical "turf" and financial stakes.

Delphi and nominal group methods have the value of encouraging independent thinking (Delbecq, 1975). Delphi seems to be particularly relevant for generating predictive information (Dalkey and Helmer, 1963) when data are poor and for resolving highly controversial issues likely to be distorted by personal interaction. Delphi, however, relies on fairly precise instructions with finite options--vague instructions are likely to induce unstable and unconfident responses (Scheibe et al., 1975). NIH-type consensus development is more flexible and permits modification as participant interaction develops. It seems appropriate for tasks such as synthesizing the state-of-the-art in a given field (Glaser, 1980). But because of the potential for domination of the group, consensus development works best where there is a substantial body of information which can be referred to.

B. APPLICABILITY OF STRUCTURED DECISION-MAKING TO RVS DEVELOPMENT

These methods are generally used in situations which require that a problem be solved--a correct answer obtained. The various participants, from their own knowledge, attempt to describe the true and potentially demonstrable state of reality. Differences between participants in such a process are based on different interpretations of the available data. Nevertheless, the objectives of the individual members are similar and coincident with the objectives of the group as a whole.

Setting relative values is a different activity where no objectively correct solution exists. RVSs, rather, could reflect preference where participants, perhaps representatives of diverse political groups, have different interests in the outcome of the group decision process.

Our review of available position statements and recommendations by organizations with some interest in the relative values of medical procedures confirm that there is no appropriate conceptual basis for construction of an RVS. As noted above, existing RVSs have been charge-based and generally reflect the views of practicing physicians--usually implicitly, by relying on analysis of physician charge data. Yet, even in the physician community there is expressed concern that charges do not adequately reflect "effort." The American Society of Internal Medicine strongly argues that reimbursement more adequately remunerate the "cognitive" functions that physicians perform. Translated into RVS development, this position logically would call for a greater value for visits relative to technical procedures.

Parties other than physicians naturally would have other goals or preferences in constructing an RVS. Government, for example, would want an RVS to help control total expenditures or to encourage broader access to physician services. The California Medicaid program, for example, implemented a de facto social preference approach to RVS setting when it implemented a uniform, statewide Medicaid fee schedule in 1976. In order to encourage greater physician participation in Medicaid and a different distribution of physicians' services, the state set up schedules which, on average, provided a 30 percent increase for maternity care services, a 20 percent increase for primary care services, and a 65 percent increase for anesthesia services. All other services received an increase of 9.5 percent. These charge increases were selected not to conform with physicians' charge profiles, but rather to accomplish broader policy objectives. A desired policy objective in fact took place--primary care physicians dramatically increased the number of Medicaid patients they treated (Holahan et al., 1981).

Other examples, likewise, demonstrate that there is no agreement on the appropriate conceptual basis for construction of an RVS. Should the RVS represent resource costs, relative prices, efficient prices, or social preferences? Large quantities of data efficiently processed would be useful but not decisive for group decision participants with differing interests. We are not currently aware of a theoretical basis for establishing a normative definition of what RVSs should represent. To the extent that various goals of an RVS would be encouraged as part of a group decision process, e.g., more adequately reward cognitive services, the process becomes less that of finding a solution and more that of achieving the most "politically" acceptable choice. The findings of research on formal, group decision-making for problem solving tasks are unlikely to be valid for group choice tasks in which participants have a stake in the outcome and no objectively correct solution exists (Wagner, 1982).

Delphi and the other formal techniques depend upon using experts to provide best guesses about the state of things for which the necessary data are not available. Much has been written about the pitfalls of expert panel selection in Delphi (Dalkey, 1969; Bedord, 1972). Experts in estimating time, skill, complexity, etc.--presumably physicians--would have a financial stake in the outcome of the RVS determinations and thereby have a substantial conflict of interest if empaneled to determine an RVS. It is necessary to recognize clearly that unless narrowly defined, RVS development on a national scale would be basically a social and political decision-making process and not a technical one.

C. GUIDELINES FOR ACHIEVING SOCIAL CHOICE CONSENSUS

Wagner has recently provided some guidelines applicable to achieving group social choice consensus where there is no correct solution (Wagner, 1982). That paper reviewed in some detail the impact of process, i.e., the structure, procedures and information needs, on committee decision-making in the allocation of health capital resources. While the full discussion is beyond the scope of this report, a few points have particular relevance to RVS determinations through group decision-making.

Three factors related to membership in the decision-making group will affect the outcomes of the group decisions: (1) group size; (2) the representational structure of the group; and (3) the level and mix of technical expertise in the group.

Group Size

Wagner finds that the larger the group, the more that cognitive and intellectual resources can be brought to bear on the decision (Ford and McLauglin, 1976); but larger groups also take longer to resolve conflicts (Filley, 1978). Larger face-to-face groups also tend to experience more tension, greater inequality of participation among group members and some loss of information (Davis, 1969). Additional members offer additional solutions, but at a decreasing rate--simulations have shown that the additional awareness brought to the group by the ninth and successive members is very low (Ford and McLaughlin, 1976). Wagner concludes that a practical limit to the size of face-to-face decision-making groups is ten (Wagner, 1982).

Representational Structure

But with RVS setting, as with resource allocation, representational structure of the group becomes crucially important. Physicians would have different interests among themselves about the relative importance placed on the factors which can be considered in establishing an RVS. Primary care physicians, as demonstrated by the position of the ASIM and its endorsement by the American Academy of Family Practitioners, would be expected to emphasize the importance of cost-related factors, especially time, in contrast to procedure-oriented specialists, who do quite well on a charge-based system.

Consumer representatives could be expected to advocate that greater weight be placed on procedures which might result in accomplishment of certain social goals, such as more equitable distribution of physicians or reduction in "unnecessary" operations. Insurers and government representatives naturally would have a primary concern about costs--not only about physician costs themselves but also about related hospital costs generated by physician behavior.

Considering these varied and strongly held positions, a basic decision on representation on an RVS decision-making group is whether members should represent particular interests or whether members represent the community at large--a "commission." With all of the interests that would demand representation, the former approach would likely result in an unwieldy group far in excess of the desirable group limit of 10 members. A commission approach probably would be workable, recognizing the potential for cooptation of commission members by interest groups.

However, it must be recognized that because it was established to represent the community at large, commission(s) generally are expected to substantiate every assertion made and to provide support for every position taken.

In order for its product, then, to stand up under public scrutiny, a commission has a tremendous need for technical staff work to support its deliberations. An example of the effort required of a "commission" in a major health policy area was the work of the Graduate Medical Education National Advisory Committee (GMENAC) in health manpower planning. (USDHHS, 1980) GMENAC was charged in 1976 as an advisory body to DHHS without regulatory powers, began its work in 1977, published an Interim Report in 1979, and a very large final report when it culminated its activities in 1980. The Committee had a Chairman, Executive Secretary and 22 members. It appointed approximately 180 physicians and thirty nonphysician health care providers to serve on formal advisory panels. A support office in DHEW included 15 professionals--statisticians, epidemiologists, and economists--who were integral members of GMENAC. The committee met monthly or bimonthly in public session. Technical panels conducted their work in separate meetings and provided frequent reports to the full GMENAC Committee.

One should anticipate that a commission charged with recommending a comprehensive system of relative values to be used as the basis for physician reimbursement under government-run programs would require substantial time and resources and would provide a focal point for public debate and disagreement.

An alternative method for fee setting, used in the Canadian provinces, is a formal negotiation between the medical association and the government. In this approach, physicians would have to resolve their particular claims internally, and government, likewise, would have to develop positions which balance the claims of health status improvement, cost concerns, equity, etc. Our preliminary review of the negotiating process in Quebec suggests that a primary basis for government's negotiating position on relative fees is to equalize rates of return to training in various specialties. Canadian provincial governments and the West German government, as other examples, have generally been more interested in adjusting the fees paid to doctors to control expenditures in the aggregate than in fine-tuning the relative values of the various procedures to achieve other policy goals.

Role of Experts

Regardless of the composition of the decision-making body, the proper role to be played by experts is an important consideration. As pointed out above, experts, whether on relatively informal committees or participating through formal Delphi or nominal group approaches, are best employed where there is a theoretically correct solution and a lack of conclusive or precise data. Expertise is a mixed blessing where decisions essentially involve social choice (Wagner, 1982). Where issues are complex, experts can be expected to do a better job in making good choices than non-experts. At the same time, those with technical expertise also represent particular constituencies. RVS "experts" would likely be drawn from the physician community and the insurance industry and could be expected to have a predisposition to support a charge-based approach with which they are familiar and comfortable.

It has been found that a group member(s) who uniquely possesses information relevant to a group's decision will make more attempts to influence the group and will have higher status in the group (Richardson et al., 1973). Experts can actually reduce the performance of groups, even on problem-solving tasks, particularly if the topic is in reality outside the sphere of expertise, if conformance to experts' opinions sets in before all opinions are expressed, or if habit leads to dependence on experts (Collins and Guetzgow, 1964).

Committee Procedures

The procedures by which the committee functions have a significant impact on the outcomes of deliberations. In operational terms, procedures include voting rules, the role of subcommittees, rules of order and agendas. While concluding that procedures definitely make a difference, Wagner feels that there are at present no general models available to allow analysis of the impact of any procedure on the outcome of complex decision-making groups (Wagner, 1982).

Majority rule is the standard voting procedure for most public and quasi-public committees. Alternative voting rules include unanimity and veto. The latter rules would likely lead to a high incidence of unresolved decisions for the overall committee but might have applicability for technical tasks carried out by small groups.

Agenda setting, defined as the scope and sequence of decision making, is another important procedural issue. In RVS development, a critical decision will be whether to use an existing RVS--presumably charge-based--as the basis for modification or whether to start *de novo* by initially setting out criteria on which to base RVS development and then seeking data. The outcomes are likely to be quite different depending on this decision.

Similarly, the precise instructions to subcommittees are likely to affect overall outcome. Subcommittees can be tightly restrained by being asked to carry out very specific technical tasks, i.e., estimating time per procedure, or they can have a shot at the broader consideration of "value" of procedures. The broader the charge to the subcommittees, the more likely that these experts, with interests in the outcome RVS setting, would influence overall committee deliberations.

Subcommittees of experts can be asked to carry out a number of functions, whether simultaneously or sequentially, i.e., making time estimates, skill/ complexity estimates, overhead expense estimates, etc., thereby facilitating administrative efficiency. It would need to be decided, however, whether a single group could justifiably function as experts in rather diverse areas. Again, the greater the number of tasks assigned to single subcommittees, the more influence the subcommittees are likely to have.

Even rules of order can be expected to affect outcome. Experiments have shown that the use of prespecified rules of order (e.g., Robert's Rules) encourages choices that lie within the majority core. (Hoffman and Plott, 1980) Rule-free decision processes tend to inhibit the search for information and enhance the power of coalitions, both of which tend to move away from the majority core. (Wagner, 1982)

In summary, research on the impact of procedures and outcomes on social choice decision-making has demonstrated unequivocally that procedures make a difference. It is clear that the procedures by which any RVS-setting committee functions should be given prompt, explicit consideration, particularly in relation to decisions on group membership and representation.

<u>Data Needs</u>

A particular problem for RVS setting is the lack of adequate data on which committee members can base their recommendations. For example, a member may wish to compensate physicians primarily on the basis of time spent on a procedure--yet, there is no data base which provides reliable time information for more than a small percentage of the several thousand currently identified procedures. There are even greater data gaps on skill, complexity and other cost factors which can be used to construct relative values. Only charge-based data is consistently and readily available. The implication is that even a commission attempting to balance various bases for RVS setting would not be able to carry out this goal because of the data gaps.

A possible solution would be the use of experts on subcommittees to provide technical information to the full committee empowered to set relative values. Many of the data needs, particularly related to costs, lend themselves to group decision-making, particularly of the Delphi type. Groups of experts appropriate to specific, clearly defined tasks, i.e., determining average operating room times for a list of surgical procedures, determining the skills required to perform a set of procedures, can be effectively used to establish the data bases that the RVS setting group would consider in its deliberations.

GMENAC provides a model for the use of expert panels to provide best guess estimates where data is scarce or outdated. Delphi panels of experts, composed of physicians from the various specialties and non-physician providers, provided revised data estimates on the incidence/prevalence of disease, on the norms of care, and on physician productivity in each specialty for use in a mathematical model of physician need. (USDHHS, 1980) There were separate panels addressing each specialty or group of closely related specialties. (McNutt, 1981)

In experimental RVS situations, expert panels have been used to fill in data gaps related to the costs of physicians' activities. For example, Hsiao and Stason (1979) employed groups of surgeons in five specialties and broadly representative of various practice modes to assess the complexity of different surgical services. They employed a modified Delphi technique to obtain estimates of relative complexity of procedures and found that the "experts" within specialties displayed little variation in their estimates despite the diverse characteristics of the practices represented.

In a nonexperimental situation, Mountain Medical Group, a preferred practice organization in Denver, used subcommittees of specialists as experts to advise a committee of only six physicians which had the overall responsibility of determining a plan-wide RVS (personal communication Manfred Oliphant, M.D.). The subcommittees were apparently able to achieve consensus on assessing factors on which they were experts, i.e., time, complexity, skill, and offered their recommendations to the full committee. In some cases, apparently, the committee modified recommendations of subcommittees based on economic considerations, e.g., the relative value of Caesarean sections were downgraded somewhat in an attempt to discourage their use.

It must be emphasized, however, that Mountain Medical used a charge-based schedule, which can be traced back ultimately to the California RVS, as the primary basis for its consideration of RVSs. In a process as difficult and complex as establishing an RVS, it is evident that an already established schedule will be given credence and relied upon to a great extent. Nevertheless, the Mountain Medical experience is a concrete demonstration of at least one way to provide data to permit modification of an existing RVS. With enough resources, a commission could start essentially from scratch by identifying the data bases it would wish to have available and, then, utilizing subcommittees of experts constituted in one of the formal group decision-making approaches to provide best guesses to fill the data gaps.

D. SUMMARY AND CONCLUSION

Unlike many other problem-solving tasks in health care, RVS setting is fundamentally a judgment-laden task for which there is no correct solution. In such situations, undue reliance should not be placed on the judgments of expert professionals who have a stake in the outcome. Formal group decision-making methods, such as Delphi, are less appropriate for judgment decisions.

At the same time, expert groups can be effectively used to provide best-guess estimates where reliable data are not available.

In establishing a group to develop a national RVS, a fundamental decision would be whether the group should directly represent the various interests with a stake in the RVS or whether the group should constitute a commission that theoretically represents the community at large. A commission is likely to consume a great deal of resources but, theoretically, would be able to adopt a broader view on what should constitute the value of a range of medical services.

The literature on social choice consensus is limited but does demonstrate that the structure and process of the group matters. At this time, there are no general models available to allow analysis of the impact of any procedure on the outcome of complex decision-making groups. At the same time, it should be recognized that each process decision, e.g., should group deliberations be based on an existing RVS?, can predictably result in a different outcome.

CHAPTER 7

SUMMARY AND RECOMMENDATIONS FOR FURTHER RESEARCH

A. SUMMARY

A relative value scale is an ordinal and cardinal ranking of physicians' services on the basis of some underlying concept of value. Once this basic ordering is established, a fee schedule is developed by applying a dollar multiplier, e.g., ten dollars per relative value unit, to procedures' relative values in order to convert them into absolute fees. If the same multiplier is used for all procedures, then relative fees are identical to relative values. However, different multipliers for different procedures (or classes of procedures) if HCFA wished to encourage or discourage certain procedures or differentially reward physicians in different specialties. In either case, a theoretically and empirically valid relative value scale is an important if not essential starting point for establishing a fee schedule.

The primary goal of this study's first year was to identify, describe, and evaluate alternative methods of constructing a relative value scale for physicians' services. This is important because, should HCFA decide to switch from C-P-R reimbursement to a fee schedule to pay for physicians' services provided to Medicare beneficiaries, then developing a relative value scale is an important intermediate step. Five basic methods were evaluated: Charge-based methods, which build relative values from data on physicians' charges; the statistical cost function approach, which would derive relative values from the parameters of a multiproduct cost function for physicians' services; time-based methods, which use data on the amount of time physicians spend performing various procedures; micro-costing and time/motion study methods, which use information obtained by detailed, on-site observation of the process of producing physicians' services; and consensus development/social preference methods, which rely on expert opinion and/or group decision making to arrive at relative values.

In addition to describing and evaluating the conceptual underpinnings of each class of methods, they were also assessed in terms of several pragmatic, implementation criteria. What data are required, are they available, and at what cost? What types of technical expertise are needed to construct and implement relative values? How well does the method assign relative values to certain "problem" procedures, such as new or rapidly changing procedures? How easy would it be to update relative values? Can specialty- or location-specific relative values be constructed? How difficult would it be for individual carriers to construct and/or implement the scale? What is the method's potential for incorporating into the scale costs and benefits that occur outside the direct patient-physician encounter?

Only two of the methods, the charge-based and consensus development/ social preference approaches, appear to be both theoretically sound and feasible to implement at reasonable cost. Although both of these sets of methods have some flaws, they are relatively minor compared to the drawbacks of the other three approaches. Econometric estimation of multiproduct cost functions would require much more detailed data on physicians' practices than are currently available. Furthermore, the statistical properties of cost functions with more than two outputs are unknown. Given the number of outputs that would need to be included for any meaningful relative value scale construction, the odds are that the parameter estimates would be highly unreliable and possibly implausible (e.g., negative marginal costs, which would imply negative relative values).

Micro-costing and time/motion studies are conceptually straightforward but are extremely expensive to perform. Existing micro-costing studies that include physician as well as other costs involved only a very few institutional settings, usually one. Thus, their results may be very site-sensitive. Furthermore, the extension of similar studies to many sites would be exorbitantly expensive.

Time-based methods depend critically on the assumption that a procedure's value, which includes costs and benefits, is a constant proportion of the amount of physicians' time required to provide the procedure. Except for narrow classes of procedures, however, this assumption does not generally hold. Substantial variations across procedures in nonphysician costs, physicians' investments in training, and the values of outcomes make the relationship between physicians' time input and procedures' values nonproportional and probably nonlinear. One could adjust observed time to account for these other factors. This would involve a substantial and expensive effort to collect information on variables that are difficult to observe, such as the value of a physician's time and the allocation of fixed costs among procedures. In fact, if such data were collected, then the adjusted time-based method would be essentially equivalent to the more general micro-costing and statistical cost function methods. Finally, there are no large data bases which contain information on physicians' time for procedures identified by CPT-4 or another standard procedure coding terminology.

Relative values constructed from charge data appear to have several desirable properties. Their construction is straight-forward. Large, computerized data bases of physicians' charges for procedures identified by their CPT-4 (or similar) codes are readily available at reasonably low cost. (For example, the cost of processing approximately 65,000,000 claims to construct

The Urban Institute's California Medicare/Medicaid Claims file was about $200,000.) Scales are highly invariant with respect to the particular distribution point (mean, median, 75th percentile, or 90th percentile) selected to represent a procedure's absolute charge. Scales constructed from different charge data bases are highly correlated. Scales appear to be stable over time. Finally, for the limited number of procedures for which data are available, the charge-based scale correlated well with the RV scale constructed by the Mountain Medical Affiliates of Denver, Colorado, and with a scale constructed from data on physicians' time per procedure. (See chapter 4)

The primary concern over using charges to construct relative values is that charges may be distorted because of uneven insurance coverage among different types of procedures, the inherently inflationary effects of C-P-R reimbursement systems used by many insurers, and the alleged noncompetitive structure of the market for physicians. Research has shown that insurance does indeed increase physicians' charges, as do C-P-R reimbursement and noncompetitive market structure. It does not follow, though, that if absolute charges are in some sense too high, that relative values constructed from charges will be seriously out of whack compared to what they would be if constructed from "undistorted" charges.

Whether some relative values are too high and others too low cannot be resolved through technical or statistical analysis alone. The review of consensus development/social preference methods showed that group decision-making can be applied to the task of constructing relative value scales. Developing a group decision-making process requires making numerous procedural decisions: group size, group composition, nature of group interactions, voting rules, the definition of value, criteria to be considered, specification of the purpose of the RV scale, range of procedures to be considered, use of experts, etc. How these procedural issues are resolved can obviously influence the outcome. Thus, HCFA's objectives (or those of whoever has responsibility for structuring the process) need to be known before procedural decisions are made.

Probably the most useful application of a consensus development process would be to review, evaluate, and adjust a relative value scale constructed from charges. Given some objective, for example, "Adjust relative values which are 'out of line' with respect to production costs, or HCFA's costs, or different specialists' incomes, or efficacy, or patients' costs (travel time, waiting time, etc.)," a panel would alter the relative values of procedures identified in accordance with the group's instructions. Such a process would obviously be highly political but may be necessary to implement any RV scale, regardless of how it is constructed.

B. RECOMMENDATIONS FOR FURTHER RESEARCH

This project's second year will be concerned with three questions important to the implementation of a relative value scale as a basis for paying physicians. Given the likelihood that physicians' charges will be a key part of any RV scale, does insurance distort charge-based relative values and by how much? How would the implementation of a RV scale as the basis for paying physicians influence physicians' receipts from Medicare (and Medicaid) and Medicare's and Medicaid's costs? Will altered relative values affect the quantities of specific services supplied by physicians?

The first question will be addressed using HCFA's prevailing charge data file, which contains information for approximately 110 procedures and 210 localities. Relative values for selected procedures will be computed for each locality. Locality-specific relative values will then be aggregated to the state level using locality-to-state population ratios as weights. With the state as the unit of analysis, we will use multivariate regression techniques to test the hypothesis that insurance coverage, measured by private (Blue Cross and commercial) insurance premiums per household and Medicaid spending for physicians' services, has a differential impact on procedures' relative values. The tentative alternative proposal is that procedures which are thought to be relatively less well insured, e.g., office visits, will have higher relative values where insurance coverage is high. Failure to reject the null hypothesis of no difference among procedures would imply that insurance distortion is insignificant.

We will address the second question by simulating the redistributive and cost implications of applying a RV scale to actual payments made by Medicare and Medicaid in California. The data base for this analysis will be The Urban Institute's California Physicians' Medicare and Medicaid Claims file. As described in detail in the Data Base Completion Report, the California Physician File contains all claims submitted by approximately 7,500 physicians during the first calendar quarter of each year from 1974 through 1978.

The simulations will be based on data from the 1978 file. The reference point for the simulations will be the amount and distribution of money actually spent by Medicare and Medicaid for reimbursements to the physicians in the file. Thus, the initial distribution of payments and initial total cost will be based on Medicare's and Medicaid's actual payments to physicians grouped by specialty, type of practice (solo or group), age of license (a special subsample of all physicians who obtained licenses between 1973 and 1978 is part of our file), practice location (large SMSA, small SMSA, non SMSA), and country of medical education (foreign or domestic). Separate distributions and totals will be computed for Medicare and Medicaid.

The simulations will consist of computing how Medicare's and Medicaid's payments would have changed under the assumptions of a constant price per relative value unit, but a different structure of relative values associated with each procedure. The total cost to Medicare and Medicaid can be held constant by picking the conversion factor--the price per relative value unit--such that total costs remain the same. Under these circumstances, the primary impact of an alternative relative value scale will be to alter the distribution of payments among physicians. For each group of physicians identified above, we will report the percentage change in Medicare and Medicaid payments (under the constraint that total changes in payments sum to zero, i.e., total program costs remain constant).

One RV scale which will be analyzed is the California RVS (CRVS).

A step-by-step description of how its implications for the Medicare program will be simulated should clarify how this task will be conducted.

- Each procedure is assigned its CRVS value. (This already exists on the file.)

- For each physician, the total number of CRVS units is computed by multiplying, for each procedure, the number of services provided to Medicare beneficiaries by its CRVS value. This will be done separately for each procedure type (medical, surgical, laboratory, and radiology).

- The total number of CRVS units will be computed for each major procedure type by summing over all physicians in the file.

- Total Medicare payments will be divided by total CRVS units for each procedure type to produce the four conversion factors which hold program payments constant.

- The conversion factors, which are measured in dollars per CRVS unit, are next applied to each physician's output of CRVS units to simulate how much each physician would have received if procedures' relative prices were identical to CRVS values.

- Distributional consequences are determined by summing total payments based on CRVS relative values for each of the groups of physicians identifed above and comparing these totals with actual Medicare payments.

Identical procedures would be followed to estimate the impact of converting Medicaid (Medi-Cal) to the same set of relative values. The implications of alternative relative value scales will be assessed by altering the values which are assigned in the first step of the simulation. We plan to use several of the RV scales computed in chapter 2, including the Mountain Medical Affiliate's scale. The computed scales can be adjusted to simulate the consequences of altering relative values for particular groups of procedures, e.g., increasing visits' values or reducing surgical procedures' relative values.

In order to be politically acceptable to physicians, it is probable that one condition an RVS would have to meet is that no group of physicians be made worse off (in terms of income) relative to the existing system. This, of course, implies that program costs would increase. In effect, the higher program costs, at least in the short run, would have to be considered part of the cost of implementing the RVS system. (This seems to have been the case in Canada, for example, where all physicians' incomes went up immediately following the conversion to fee schedules (Hadley et al., 1979).) Thus, the second major output of the simulations will be to compute the change in the value of the conversion factor which would meet the condition of political acceptability to physicians.

These simulation outputs will be summarized by a series of tables showing the changes in the distribution of payments for each of the groupings of physicians, the increase in program costs required to at least maintain all physicians' current payment levels, and the values of the conversion factors for each procedure type. Table 7.1 illustrates what the table for the impact on different specialties might look like. Similar tables would be produced for other distributional groupings.

Table 7.1

Impact of Alternative RVS on Payments
by Specialty-Medicare
(California, 1978)

Specialty	Current Distribution	CRVS	RVS #2	RVS #3	...	RVS #n
	Percentage Distribution of Payments					
General & Family Practice						
Internal Medicine						
etc.						
All Specialties	100%	100%	100%	100%	...	100%
Conversion Factors - $s per RVU						
(No Change in Program Costs)						
Medical Procedures						
Surgical Procedures						
Laboratory Procedrues						
Radiological Procedures						
Conversion Factors - $s per RVU						
(No Decreases in Payments to Physicians						
Medical Procedures						
Surgical Procedrues						
Laboratory Procedures						
Radiological Procedures						
Total Cost Increase - $					...	

A key assumption underlying this first set of simulations is that the number of times a procedure is performed is not influenced by the relative value scale. In other words, the aggregate (and individual) physicians' supply functions are assumed to be vertical, i.e., independent of price. Previous research done at The Urban Institute and elsewhere suggests that this is not the case. First, it is reasonably well-known that the number of physicians treating Medicaid and Medicare patients goes up as Medicare and Medicaid unit prices increase relative to billed charges. Second, it also appears that the Medicaid and Medicare patient loads per physician also increase (Hadley, 1979; Paringer, 1980). Finally, although procedure-specific supply functions have not been estimated, it seems probable that the provision of at least some procedures, presumably those that have close medical substitutes or are of marginal therapeutic value, are sensitive to the payment received.

Under these circumstances, it seems that the results of the constant quantity simulations should be subject to a "sensitivity" analysis based on alternative assumptions about the elasticities of the implicit procedure-specific supply functions. (These assumptions could be made arbitrarily, built from best-guess estimates from prior research, or based on the direct estimation of simple supply functions using the California physician data.)

We propose to explore the estimation of procedure-specific supply functions for selected procedures reported in our California Physicians' Medicare-Medicaid Claims file. Procedures will be selected for which there are reasonably close alternatives that can be provided by the same physician. Two obvious sets of procedures are office visits and hospital visits of various durations. The question we would address, for example, is whether the number of brief office visits is influenced by the amount paid for a limited office visit. If it is, then this would suggest that manipulating relative values, and ultimately relative fees, would alter the mix of services physicians provide.

A second approach, though more tentative at this time, would be to obtain data from Canada's provincial health plans on expenditures, quantities, and fees for specific services. If such information is available from published sources for several years, then an aggregate, time-series, cross-section data base could be constructed to analyze whether variations in relative fees influence the supplies of various classes of services. One advantage of using Canadian data is that all provinces offer complete coverage for all services to their total population. Thus, the effects of variations in insurance coverage among procedures and population groups would not be present.

The information presented in this report demonstrates that credible RV scales are feasible to construct at relatively low cost. In the project's second year we plan to focus on the perhaps more important question of how relative fees influence physicians' behavior.

BIBLIOGRAPHY

Averch, H. and L.L. Johnson. "Behavior of the Firm Under Regulatory Constraint." American Economic Review, 52 (December 1962):1052-1069.

Barnes, Ralph M. Motion and Time Study. New York: John Wiley & Sons, Inc., 1958.

Bedford, Michael T. "The Value of Competing Panels of Experts and the Impact of 'Drop-outs' on Delphi Results." Delphi: The Bell Canada Experience, Bell Canada, October 1972.

Blalock, H. Social Statistics. New York: McGraw-Hill, Inc., 1972.

Brouhard, Chris E. "System Monitors Productivity in Diagnostic Imaging." Hospitals 55(4) (July 1981):167-175.

Brown, R.; D. Caves; and L. Christensen. "Modelling the Structure of Cost and Production for Multi-Product Firms." Southern Economic Journal (July 1979):256-267.

Bunning, R.L. "The Delphi Technique: A Projection Tool for Serious Inquiry," in The 1979 Annual Handbook for Group Facilitators, J.E. Jones, J.W. Pfeiffer (eds.), La Jolla, Calif.: University Associates, 1979.

Campbell, R.M. "The Delphi Technique: Implementation in the Corporate Environment." Management Services 5 (November-December 1978):37-42.

Carlson, Cathy. "Data Base Completion Report." Working Paper No. 3075-03, The Urban Institute, May 1982.

Caves, D.W.; L.R. Christensen; and J.A. Swanson. "Productivity Growth, Scale Economies, and Capacity Utilization in U.S. Railroads, 1955-1974." American Economic Review (December 1981):994-1002.

Christensen, L.R.; D.C. Christensen; and P.E. Schoech. "Econometric Estimation of Scale Economies in Telecommunications." Econometric Analysis of Telecommunications: Theory and Applications, North-Holland Publishing Co., 1982.

College of American Pathologists, "Laboratory Workload Recording Method," 1981.

Collins, B.E.; and H. Guetzkow. A Social Psychology of Group Processes for Decision-Making. New York: John Wiley & Sons, Inc., 1964.

Dalkey, N.C. Experiments in Group Prediction. The Rand Corporation, P-3820, March 1968.

Dalkey, N.C. The Delphi Method: An Experimental Study of Group Opinion. The Rand Corporation, RM-5888-PR, June 1969.

Dalkey, N.C.; and O. Helmer. "An Experimental Application of the Delphi Method to the Use of Experts." Management Science 9(3) (1963):458.

Davis, James H. Group Performance. Addison-Wesley Publishing Company, 1969.

Delbecq, A., et al. Group Techniques for Program Planning, Glenview, Ill.: Scott, Foresman & Co., 1975.

Eisenberg, Barry S. "Information Exchange Among Competitors: The Issue of Relative Value Scales for Physicians' Services." Journal of Law and Economics 23 (October 1980):441-460.

Farquhar, J.A. A Preliminary Inquiry into the Software Estimation Process. The Rand Corporation, RM-6271-PR, August 1970.

Filley, A.C. "Committee Management: Guidelines from Social Science Research." California Management Review 13(1) (Fall 1970):13-21.

Filley, A.C. "Some Normative Issues in Conflict Management." California Management Review 21 (Winter 1978):61-66.

Fineberg, H.V. "Clinical Chemistries: The High Cost of Low-Cost Diagnostic Tests," in Medical Technology: The Cluprit Behind Health Care Costs. The Proceedings of the 1977 Sun Valley Forum on National Health, DHEW Pub. no. (PHS) 79-3216, 1978.

Ford, R.C.; and F. McLaughlin. "Effects of Group Composition on Decision Search Activity." Journal of Business Research (1) (1976):15-24.

Glaser, E.M. "Using Behavioral Science Strategies for Defining the State of the Art." Journal of Applied Behavioral Science, 16 (1980):79.

Glenn, D.W. "Laboratory Cost-Accounting and Workload Analysis." Chapter 27 in Administration and Supervision in Laboratory Medicare. J.R. Snyder (ed.) Philadelphia: Lippincott/Harper, 1982.

Greene, W.H.; and L.R. Christensen. "Economies of Scale in U.S. Electric Power Generation." Journal of Political Economy, Vol. 84, No. 4, P.I. (August 1976):655-76.

Hadley, Jack. "An Econometric Analysis of Physician Participation in the Medicaid Program." Working Paper No. 998-9, The Urban Institute, April 1978.

Hadley, Jack. "Physician Participation in Medicaid: Evidence from California." Health Services Research 14 (Winter 1979):266-80.

Hadley, Jack. "Financial Incentives and Physician Distribution: Evidence from Canada and the United States." Working Paper No. 1306-02-07, The Urban Institute, July 1982.

Hadley, Jack, et al. "Can Fee-For-Service Reimbursement Coexist with Demand Creation?". Inquiry (Fall 1979):247-58.

Hadley, Jack, et al. "Preliminary Research Plan for Alternative Methods for Developing a Relative Value Scale of Physicians' Fees." The Urban Institute (November 1981):8-12.

Harris, T.R. "A Poor Measure of Physicians' Service." The Internist (June 1977).

Hoffman, E. and C. Plott. "Pre-Meeting Discussions and Possibility of Coalition-Breaking Procedures in Majority Rule Committees." California Institute of Technology, Social Science Working Paper No. 281, 1980.

Holahan, J.; Janet Gornick; and Daniel Nichols. "Medicaid Reimbursement for Physician Services." Working Paper No. 1298-08, The Urban Institute, 1981.

Holahan, J.; and W. Scanlon. Price Controls, Physician Fees and Physician Incomes. Washington, D.C.: The Urban Institute, 1978.

Holahan, John, et al. "Medicaid Fee Controls and Physician Behavior: Preliminary Evidence from California." Working Paper No. 1250-03, The Urban Institute, March 1981.

Hsiao, William and Willian Stason. "Toward Developing a Relative Value Scale for Medical and Surgical Services." Health Care Financing Review (Fall 1979):23-28.

Kirchener, Merian. "The Pros and Cons of 'A La Carte' Fees." Medical Economics (June 1979):1-4.

Lee, Robert and Jack Hadley. "Supplying Physicians' Services to Public Medical Care Programs." Working Paper No. 1145-17, The Urban Institute, 1979.

Lee, Robert and Jack Hadley. "Physicians' Fees and Public Medical Care Programs." Health Services Research (Summer 1981).

MacDonald, L. and L.F. Reuter. "A Patient-Specific Approach to Hospital Cost Accounting." Health Services Research (Summer 1973).

Markel, Gene. "Hospital Utilization Effects of Care Reimbursement for Medical Care." Physicians and Financial Incentives. HCFA Conference Proceedings, December 1980.

McNutt, David R. "GNEMAC: Its Manpower Forecasting Framework." American Journal of Public Health (October 1981):1116-1124.

Paringer, Lynn. "Medicare Assignment Rates of Physicians: Their Responses to Changes in Reimbursement Policy." Health Care Financing Review (Winter 1980):75-90.

Paringer, Lynn. "The Medicare Economic Index: Impact on Program Costs and Beneficiary Liability." Working Paper No. 1306-01-03, The Urban Institute, June 1981.

Paringer, Lynn. "The Distributional Effects of the Medicare Economic Index on Physicians and Beneficiaries." Working Paper No. 1306-01-05, The Urban Institute, July 1981a.

Paringer, Lynn. "The Effect of the Medicare Economic Index on Reasonable Fees: Evidence from California." Working Paper No. 1306-01-04, The Urban Institute, July 1981b.

Penner, D.W. "Changes in the Workload Recording Method? What They Mean and How to Cope." Pathologist 36:1 (January 1982) :27-29.

Perry, S. and J.T. Kalberer, Jr. "The NIH Consensus--Development Program and the Assessment of Health Technologies: The First Two Years." New England Journal of Medicine, 303 (1980):169.

Richardson, R.W. "Comparison of Cost of Preparing Reagents in Laboratory with Cost of Using Commercial Kits." The Lancet (December 1977).

Richardson, R.W., et al. "Expert Power: A Behavioral Interpretation." Sociometry 36 (1973).

Sheibe, M.; M. Skutsch; and J. Schofer. "Experiments in Delphi Methodology." The Delphi Method: Techniques and Applications, Reading, Mass.: Addison-Wesley Publishing Company, 1975.

Showstack et al. "Fee-for-Service Physician Payment: Analysis of Current Methods and Their Development." Inquiry 16 (Fall 1979):230-246.

Shuman, L.; H. Wolfe; and M. Perlman. A Generalized Model for Hospital Microcosting, 1971.

Shuman, L., H. Wolfe and M. Perlman. "Model for Hospital Microcosting." Industrial Engineering (July 1973):39-43.

Shuman, L.J., et al. Manuel for Implementing Radiology Micro-Costing. Department of Industrial Engineering, Systems Management Engineering and Operations Research, University of Pittsburg, August 1972.

Sloan, F.; J. Cromwell; and J. Mitchell. "Physician Participation in State Medicaid Programs." Journal of Human Resources 13 (Supplement 1978): 211-245.

Sloan, F.; J. Cromwell; and J. Mitchell. "Private Physicians and Public Programs." Lexington, MA: Lexington Books, 1978.

BIBLIOGRAPHY (Continued)

Sobaski, William. "Effects of the 1969 California Relative Value Studies on Costs of Physician Services Under SMI." DHEW Pub. No. SSA 75-11702, June 1975.

Stilwell, A. "Costs of a Clinical Chemistry Laboratory." Journal of Clinical Pathology, 34 (June 1981):589-594.

U.S. Congress, Office of Technology Assessment, Strategies for Medical Technology Assessment, Washington, D.C., September 1982.

U.S. Department of Health and Human Services, Summary Report of the Graduate Medical Education National Advisory Committee, DHHS Publication No. (HRA) 81-651, Washington, D.C., September 1980.

Wagner, Judy. "The Committee Resource Allocation Process: Implications for Health Care Capital Expenditure Limits." Final Report, The Urban Institute, 1982.

Wild, L.K. "Bacteriology Laboratory Costs: A Universally Applicable Model," Mathematics Programs for Activity Analysis. Paul Van Moeske (ed.), New York: North Holland/American, 1974.

APPENDIX TO CHAPTER 2

Selected Relative Value Scales: High-Volume Medicare Procedures

APPENDIX TO CHAPTER 2

Selected Relative Value Scales:
High-Volume Medicare Procedures

In this appendix are the four relative value scales which were the focus of section D. In that section, we reported the results of correlation analyses of RVSs constructed from charge data from five sources; the HCFA Prevailing Charge file, the Health Insurance Association of America's (HIAA's) surgical charge file, two years of data from The Urban Institute's samples of California Medicare claims--1974 and 1978--and the relative value schedule developed by Mountain Medical Affiliates (MMA) of Colorado. For a set of high-volume Medicare procedures,* we identified "representative" charges from the several files as follows:

HCFA -- 75th percentiles of distributions of regional median customary fees

HIAA -- 75th percentiles of distributions of regional median customary charges

UI California -- 75th percentiles of distributions of local median billings

MMA -- The relative (dollar) value of each procedure as reported by MMA Inc.

The accompanying table contains relative values for these procedures defined as the ratio of each procedures' representative charge and the representative charge for the numeraire procedure--Needle Puncture of Bursa. Each procedure is identified by its HCFA Prevailing File "Line Number," its California Relative Value Studies code, and an equivalent CPT-4 code number, as well as by a simple descriptive name.

---

*These are the 103 high-volume procedures of the HCFA prevailing Charge file. Not all 103 are represented in each of the five data sources; hence, not all procedures are represented on each RVS.

Besides the five RVSs discussed in Section 2.D, the appendix contains four additional scales--constructed from HCFA, HIAA, and both UI California data files. We constructed these scales via the algorithm just described with one difference; in the four new scales, distributions' means rather than 75th percentile points are the "representative" charge values.

SUMMARY OF HCFA,MMA,HIAA AND CAL REL VALUE SCALES, PROCEDURES SORTED BY HCFA MEAN 50TH PCTL SCALE

	HCFA		MMA	HIAA		CAL78		CAL74	
DIST POINT	MEAN 50TH	75TH 50TH	MMA	MEAN 50TH	75TH 50TH	MEAN BA	75TH BA	MEAN BA	75TH BA
HEMOGLOBIN, COLORIMETRIC									
PROC#: 85018 CPT-4: 85018 LINENO: 82									
REL VALUE	0.12	0.13	0.20	.	.	0.19	0.19	0.19	0.17
SCALE RANK	1.0	2.0	4.5	.	.	3.0	2.5	3.0	3.0
WHITE BLOOD CELL COUNT									
PROC#: 85048 CPT-4: 85048 LINENO: 84									
REL VALUE	0.12	0.13	0.20	.	.	0.22	0.21	0.20	0.17
SCALE RANK	2.0	2.0	4.5	.	.	4.0	4.0	4.0	3.0
URINALYSIS, CHEMICAL									
PROC#: 81005 CPT-4: 81005 LINENO: 99									
REL VALUE	0.13	0.17	0.08	.	.	0.17	0.19	0.16	0.15
SCALE RANK	3.0	5.0	1.5	.	.	1.0	1.0	1.0	1.0
HEMATOCRIT									
PROC#: 85014 CPT-4: 85014 LINENO: 88									
REL VALUE	0.13	0.13	0.20	.	.	0.19	0.19	0.19	0.17
SCALE RANK	4.0	2.0	4.5	.	.	2.0	2.5	2.0	3.0
FECES, OCULT BLOOD, SCREENING									
PROC#: 89205 CPT-4: 82270 LINENO: 96									
REL VALUE	0.13	0.17	0.08	.	.	0.23	0.23	0.22	0.25
SCALE RANK	5.0	5.0	1.5	.	.	6.0	5.5	5.0	6.5
COLLECT & HANDLE SPEC'S									
PROC#: 99060 CPT-4: 99000 LINENO: 39									
REL VALUE	0.16	0.17	0.50	.	.	.	.	.	.
SCALE RANK	6.0	5.0	20.0	.	.	.	.	.	.
URINALYSIS, COMPLETE									
PROC#: 81000 CPT-4: 81000 LINENO: 98									
REL VALUE	0.17	0.20	0.25	.	.	0.23	0.23	0.23	0.21
SCALE RANK	7.0	7.0	7.5	.	.	5.0	5.5	6.0	5.0
SEDIMENTATION RATE									
PROC#: 85650 CPT-4: 85650 LINENO: 92									
REL VALUE	0.18	0.23	0.25	.	.	0.27	0.27	.	.
SCALE RANK	8.0	8.0	7.5	.	.	7.0	7.0	.	.

DIST POINT	HCFA MEAN 50TH	HCFA 75TH 50TH	MPA MPA	HIAA MEAN 50TH	HIAA 75TH 50TH	CAL78 MEAN BA	CAL78 75TH BA	CAL74 MEAN BA	CAL74 75TH BA
FLOCCULATION TEST									
PROC#: 86592 CPT-4: 86592 LINENO: 87									
REL VALUE	0.21	0.27	0.37	.	.	0.31	0.36	0.32	0.30
SCALE RANK	9.0	10.5	11.5	.	.	9.0	12.0	9.0	8.5
PROTHROMBIN TIME TEST									
PROC#: 85610 CPT-4: 85610 LINENO: 91									
REL VALUE	0.22	0.27	0.41	.	.	0.34	0.34	0.35	0.31
SCALE RANK	10.0	10.5	15.0	.	.	11.0	8.0	10.0	10.0
REES-ECKER PLATELET COUNT									
PROC#: 85580 CPT-4: 85580 LINENO: 89									
REL VALUE	0.22	0.27	0.20	.	.	0.35	0.35	0.32	0.30
SCALE RANK	11.0	10.5	4.5	.	.	12.0	9.5	8.0	8.5
BLOOD SUGAR									
PROC#: 82947 CPT-4: 82947 LINENO: 93									
REL VALUE	0.22	0.27	0.41	.	.	0.35	0.37	0.39	0.36
SCALE RANK	12.0	10.5	15.0	.	.	13.0	13.0	12.0	12.0
MIN OV EPAT									
PROC#: 90030 CPT-4: 90030 LINENO: 5									
REL VALUE	0.23	0.33	0.48	.	.	0.40	0.38	0.39	0.35
SCALE RANK	13.0	18.0	18.0	.	.	19.0	16.5	13.0	11.0
	284	284		.	.	985	985	867	867
BUN-UREA NITROGEN									
PROC#: 84520 CPT-4: 84520 LINENO: 94									
REL VALUE	0.24	0.30	0.57	.	.	0.36	0.38	0.40	0.38
SCALE RANK	14.0	14.5	21.5	.	.	15.0	16.5	14.0	14.5
CHOLESTEROL BLOOD TEST									
PROC#: 82465 CPT-4: 82465 LINENO: 86									
REL VALUE	0.24	0.30	0.41	.	.	0.36	0.38	0.40	0.38
SCALE RANK	15.0	14.5	15.0	.	.	16.0	16.5	16.0	14.5
URIC ACID									
PROC#: 84550 CPT-4: 84550 LINENO: 95									
REL VALUE	0.25	0.30	0.57	.	.	0.36	0.38	0.40	0.38
SCALE RANK	16.0	14.5	21.5	.	.	14.0	16.5	15.0	14.5

DIST POINT	HCFA MEAN 50TH	HCFA 75TH 50TH	HMA MMA	HIAA MEAN 50TH	HIAA 75TH 50TH	CAL78 MEAN BA	CAL78 75TH BA	CAL74 MEAN BA	CAL74 75TH BA
POTASSIUM, BLOOD									
PROC#: 84132 CPT-4: 84132 LINENO: 90									
REL VALUE	0.25	0.30	0.41	.	.	0.31	0.35	0.38	0.38
SCALE RANK	17.0	14.5	15.0	.	.	10.0	11.0	11.0	14.5
PAP TEST									
PROC#: 88150 CPT-4: 88150 LINENO: 97									
REL VALUE	0.27	0.33	0.37	.	.	0.41	0.38	0.45	0.50
SCALE RANK	18.0	18.0	11.5	.	.	20.0	16.5	19.0	19.0
AUTOMATED BLOOD COUNT									
PROC#: 85021 CPT-4: 85021 LINENO: 83									
REL VALUE	0.28	0.33	0.29	.	.	0.30	0.35	0.24	0.25
SCALE RANK	19.0	18.0	10.0	.	.	8.0	9.5	7.0	6.5
COMPLETE BLOOD COUNT									
PROC#: 85031 CPT-4: 85031 LINENO: 85									
REL VALUE	0.31	0.35	0.41	.	.	0.40	0.40	0.43	0.40
SCALE RANK	20.0	20.0	15.0	.	.	18.0	20.0	18.0	17.5
CULTURE, OTHER THAN BLOOD									
PROC#: 87081 CPT-4: 87081 LINENO: 81									
REL VALUE	0.35	0.40	0.49	.	.	0.38	0.38	0.41	0.40
SCALE RANK	21.0	21.0	19.0	.	.	17.0	16.5	17.0	17.5
MANIPULATION OF SPINE & GV									
PROC#: 97300 CPT-4: 97300 LINENO: 31									
REL VALUE	0.38	0.47	.	.	.	.	.	.	.
SCALE RANK	22.0	22.0	.	.	.	.	.	.	.
EKG INTERP & REPORT ONLY									
PROC#: 93010 CPT-4: 93010 LINENO: 35									
REL VALUE	0.38	0.50	0.90	.	.	0.64	0.64	0.67	0.60
SCALE RANK	23.0	23.0	27.0	.	.	22.0	21.0	22.0	21.5

DIST POINT	MCFA MEAN 50TH	MCFA 75TH 50TH	MMA BMA	HIAA MEAN 50TH	HIAA 75TH 50TH	CAL78 MEAN BA	CAL78 75TH BA	CAL74 MEAN BA	CAL74 75TH BA
BRIEF OV EPAT									
PROC#: 90040 CPT-4: 90040 LINENO: 6									
REL VALUE	0.46	0.57	0.72	.	.	0.67	0.67	0.60	0.54
SCALE RANK	24.0	24.0	24.0	.	.	24.0	23.0	20.0	20.0
BRIEF HV EPAT									
PROC#: 90240 CPT-4: 90240 LINENO: 19									
REL VALUE	0.51	0.67	0.84	.	.	0.81	0.80	0.69	0.63
SCALE RANK	25.0	27.0	26.0	.	.	27.0	27.0	24.0	23.0
LTD OV EPAT									
PROC#: 90050 CPT-4: 90050 LINENO: 7									
REL VALUE	0.51	0.67	0.96	.	.	0.79	0.77	0.72	0.65
SCALE RANK	26.0	27.0	29.0	.	.	26.0	25.5	25.0	24.0
BRIEF NHOME V EPAT									
PROC#: 90440 CPT-4: 90440 LINENO: 15									
REL VALUE	0.53	0.67	1.20	.	.	.	.	.	.
SCALE RANK	27.0	27.0	35.5	.	.	.	.	.	.
TONOMETRY									
PROC#: 92100 CPT-4: 92100 LINENO: 66									
REL VALUE	0.53	0.67	2.16	.	.	0.64	0.66	0.65	0.60
SCALE RANK	28.0	27.0	59.0	.	.	21.0	22.0	21.0	21.5
PANEL TEST 3-4									
PROC#: 80104 CPT-4: 80003 LINENO: 79									
REL VALUE	0.55	0.87	.	.	.	0.99	1.10	1.01	1.08
SCALE RANK	29.0	40.5	.	.	.	33.0	35.0	33.0	36.0
CHEMOTHERAPY FOR MALIG DIS									
PROC#: 96030 CPT-4: 90790 LINENO: 38									
REL VALUE	0.56	0.68	1.01	.	.	0.66	0.69	0.83	0.79
SCALE RANK	30.0	30.0	31.0	.	.	23.0	24.0	26.0	26.0
INTERMED OV EPAT									
PROC#: 90060 CPT-4: 90060 LINENO: 8									

	HCFA		MMA	HIAA		CAL78		CAL74	
DIST POINT	MEAN 50TH	75TH 50TH	MMA	MEAN 50TH	75TH 50TH	MEAN BA	75TH BA	MEAN BA	75TH BA
REL VALUE	0.57	0.67	1.20	.	.	0.97	0.96	0.92	0.82
SCALE RANK	31.0	27.0	35.5	.	.	31.0	30.5	30.0	30.0
BRIEF ECF V EPAT									
PROC#: 90340 CPT-4: 90340 LINENO: 14									
REL VALUE	0.59	0.80	1.20	.	.	0.91	0.96	0.86	0.80
SCALE RANK	32.0	31.0	35.5	.	.	29.0	30.5	27.0	27.5
DEBRIDEMENT OF NAILS <=5									
PROC#: 11700 CPT-4: 11700 LINENO: 40									
REL VALUE	0.61	0.83	0.75	0.74	0.72	.	.	.	.
SCALE RANK	33.0	35.0	25.0	1.0	1.0	.	.	.	.
LTD HV EPAT									
PROC#: 90250 CPT-4: 90250 LINENO: 20									
REL VALUE	0.63	0.83	1.20	.	.	0.98	0.96	0.89	0.80
SCALE RANK	34.0	35.0	35.5	.	.	32.0	30.5	28.0	27.5
RADIATION THERAPY LOW VOLT									
PROC#: 77030 CPT-4: . LINENO: 74									
REL VALUE	0.64	1.03	.	.	.	1.31	1.44	1.14	1.05
SCALE RANK	35.0	53.0	.	.	.	45.0	49.0	36.0	35.0
BREIF HOME V EPAT									
PROC#: 90140 CPT-4: 90140 LINENO: 11									
REL VALUE	0.64	0.83	1.20	.	.	1.13	1.15	1.05	1.00
SCALE RANK	36.0	35.0	35.5	.	.	38.0	37.0	34.0	33.0
BRIEF OV NPAT									
PROC#: 90000 CPT-4: 90000 LINENO: 1									
REL VALUE	0.66	0.87	1.20	.	.	0.99	0.96	0.97	0.85
SCALE RANK	37.0	40.5	35.5	.	.	34.0	30.5	31.0	31.0
CHEST X-RAY 1 VIEW									
PROC#: 71010 CPT-4: 71010 LINENO: 68									
REL VALUE	0.67	0.83	1.06	.	.	0.88	0.92	0.92	0.81
SCALE RANK	38.0	35.0	32.0	.	.	28.0	28.0	29.0	29.0

DIST POINT	HCFA MEAN 50TH	HCFA 75TH 50TH	MMA MMA	MIAA MEAN 50TH	MIAA 75TH 50TH	CAL78 MEAN BA	CAL78 75TH BA	CAL74 MEAN BA	CAL74 75TH BA
INITIAL PHYSIOTHERAPY									
PROC#: 97100 CPT-4: 97100 LINENO: 32									
REL VALUE	0.68	1.17	1.32	.	.	0.70	0.77	0.67	0.73
SCALE RANK	39.0	57.5	40.0	.	.	25.0	25.5	23.0	25.0
BRIEF EP V NPAT									
PROC#: 90500 CPT-4: 90500 LINENO: 23									
REL VALUE	0.68	0.90	0.60	.	.	1.25	1.31	1.17	1.10
SCALE RANK	40.0	42.0	23.0	.	.	42.0	43.0	38.0	37.0
PANEL TESTS 19+									
PROC#: 80119 CPT-4: 80119 LINENO: 80									
REL VALUE	0.69	1.00	0.94	.	.	.	.	.	.
SCALE RANK	41.0	48.5	28.0	.	.	.	.	.	.
LTD HOME V EPAT									
PROC#: 90150 CPT-4: 90150 LINENO: 12									
REL VALUE	0.73	0.83	1.80	.	.	1.33	1.35	1.27	1.25
SCALE RANK	42.0	35.0	51.5	.	.	47.0	45.0	43.0	44.5
ARTHROCENTESIS									
PROC#: 20610 CPT-4: 20610 LINENO: 45									
REL VALUE	0.73	0.90	1.50	1.03	1.00	1.11	1.15	1.16	1.13
SCALE RANK	43.0	43.0	41.5	4.0	3.5	37.0	37.0	37.0	39.5
EXTENDED OV EPAT									
PROC#: 90070 CPT-4: 90070 LINENO: 9									
REL VALUE	0.74	1.00	1.80	.	.	1.29	1.35	1.33	1.25
SCALE RANK	44.0	48.5	51.5	.	.	44.0	45.0	45.0	44.5
INTERMED HOME V EPAT									
PROC#: 90160 CPT-4: 90160 LINENO: 13									
REL VALUE	0.75	0.83	2.10	.	.	1.47	1.43	1.45	1.27
SCALE RANK	45.0	35.0	57.0	.	.	49.0	48.0	47.0	47.0
INTERMED HV EPAT									
PROC#: 90260 CPT-4: 90260 LINENO: 21									
REL VALUE	0.75	1.00	1.80	.	.	1.20	1.17	1.21	1.20

	HCFA		MMA	HIAA		CAL78		CAL74	
DIST POINT	MEAN 50TH	75TH 50TH	MMA	MEAN 50TH	75TH 50TH	MEAN BA	75TH BA	MEAN BA	75TH BA
SCALE RANK	46.0	48.5	51.5	.	.	39.0	39.0	40.0	41.0
LTD OV NPAT									
PROC#: 90010 CPT-4: 90010 LINENO: 2									
REL VALUE	0.76	1.00	1.80	.	.	1.27	1.31	1.28	1.25
SCALE RANK	47.0	48.5	51.5	.	.	43.0	42.0	44.0	44.5
	386	386		.	.	1640	1640	1482	1482
SURGICAL PATHOLOGY, 1-3 SPEC									
PROC#: 88302 CPT-4: 88302 LINENO: 100									
REL VALUE	0.78	0.86	1.68	.	.	0.95	1.01	1.10	1.00
SCALE RANK	48.0	39.0	45.0	.	.	30.0	34.0	35.0	33.0
EKG W/INTERP & REPORT									
PROC#: 93000 CPT-4: 93000 LINENO: 34									
REL VALUE	0.80	0.97	1.80	.	.	1.22	1.15	1.27	1.13
SCALE RANK	49.0	44.0	51.5	.	.	40.0	37.0	42.0	39.5
LTD ER V NPAT									
PROC#: 90510 CPT-4: 90510 LINENO: 24									
REL VALUE	0.80	1.00	1.80	.	.	1.49	1.46	1.52	1.40
SCALE RANK	50.0	48.5	51.5	.	.	51.0	50.5	52.0	52.0
DIALATION OF URETHRA (MALE)									
PROC#: 53600 CPT-4: 53600 LINENO: 60									
REL VALUE	0.82	0.83	1.25	0.82	0.80	.	.	.	.
SCALE RANK	51.0	35.0	39.0	2.0	2.0	.	.	.	.
SPIROMETRY									
PROC#: 94010 CPT-4: 94010 LINENO: 36									
REL VALUE	0.83	1.00	2.40	.	.	1.09	1.28	1.18	1.23
SCALE RANK	52.0	48.5	61.5	.	.	36.0	41.0	39.0	42.0
ELECTROSURG DESTRUCTION >2 LES									
PROC#: 17102 CPT-4: 17102 LINENO: 42									
REL VALUE	0.84	1.00	0.25	.	.	.	.	.	.
SCALE RANK	53.0	48.5	9.0	.	.	.	.	.	.

DIST POINT	HCFA MEAN.50TH	HCFA 75TH.50TH	MMA MMA	HIAA MEAN.50TH	HIAA 75TH.50TH	CAL78 MEAN.BA	CAL78 75TH.BA	CAL74 MEAN.BA	CAL74 75TH.BA
PACEMAKER. ELECTRONIC MONITORI									
PROC#: 93795 CPT-4: 93795 LINENO: 101									
REL VALUE	0.90	1.13	1.92	.	.	.	.	.	.
SCALE RANK	54.0	54.5	56.0	.	.	.	.	.	.
COMPLETE EYE EXAM NPAT									
PROC#: 92000 CPT-4: 92002 LINENO: 64									
REL VALUE	0.91	1.17	1.68	.	.	1.34	1.38	1.45	1.38
SCALE RANK	55.0	57.5	46.0	.	.	48.0	47.0	48.0	51.0
INTERMED OV NPAT									
PROC#: 90015 CPT-4: 90015 LINENO: 3									
REL VALUE	0.93	1.33	3.00	.	.	1.66	1.73	1.83	1.80
SCALE RANK	56.0	62.5	65.5	.	.	55.0	56.0	57.0	59.0
CHEST X-RAY 2 VIEWS									
PROC#: 71020 CPT-4: 71020 LINENO: 69									
REL VALUE	0.94	1.13	1.59	.	.	1.22	1.19	1.25	1.11
SCALE RANK	57.0	54.5	43.5	.	.	41.0	40.0	41.0	38.0
PSYCHOTHERAPY 25-30 MIN									
PROC#: 90606 CPT-4: 90843 LINENO: 30									
REL VALUE	0.94	1.17	.	.	.	1.33	1.35	1.39	1.25
SCALE RANK	58.0	57.5	.	.	.	46.0	45.0	46.0	44.5
RADIATION THERAPY SUPER VOLT									
PROC#: 77040 CPT-4: 77400 LINENO: 75									
REL VALUE	0.97	1.47	2.38	.	.	1.69	1.80	1.64	1.63
SCALE RANK	59.0	64.5	60.0	.	.	57.0	57.0	53.0	54.0
COMP EYE EXAM W/REFRACT.									
PROC#: 92001 CPT-4: . LINENO: 65									
REL VALUE	0.98	1.17	.	.	.	1.68	1.71	1.84	1.69
SCALE RANK	60.0	57.5	.	.	.	56.0	55.0	58.0	55.0
EXTENDED HV EPAT									
PROC#: 90270 CPT-4: 90270 LINENO: 22									
REL VALUE	1.00	1.33	2.40	.	.	1.65	1.64	1.80	1.75
SCALE RANK	61.0	62.5	61.5	.	.	54.0	54.0	55.0	57.5

DIST POINT	HCFA MEAN 50TH	HCFA 75TH 50TH	MMA MMA	HIAA MEAN 50TH	HIAA 75TH 50TH	CAL78 MEAN BA	CAL78 75TH BA	CAL74 MEAN BA	CAL74 75TH BA
NEEDLE PUNCTURE, BURSA NUM									
PROC#: 20605 CPT-4: 20605 LINENO: 48									
REL VALUE	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
SCALE RANK	62.0	48.5	30.0	3.0	3.5	35.0	33.0	32.0	33.0
INTERMED FR V NPAT									
PROC#: 90515 CPT-4: 90515 LINENO: 25									
REL VALUE	1.03	1.60	3.00	.	.	2.15	2.25	2.13	2.00
SCALE RANK	63.0	66.0	65.5	.	.	62.0	63.0	61.0	61.0
RADIATION THERAPY MEGA VOLT									
PROC#: 77050 CPT-4: . LINENO: 76									
REL VALUE	1.05	1.47	.	.	.	2.24	2.47	2.79	2.25
SCALE RANK	64.0	64.5	.	.	.	64.0	65.0	68.0	65.5
HIP X-RAY, COMPLETE									
PROC#: 73510 CPT-4: 73510 LINENO: 71									
REL VALUE	1.06	1.32	1.59	.	.	1.47	1.46	1.48	1.33
SCALE RANK	65.0	61.0	43.5	.	.	50.0	50.5	49.0	49.0
BIOPSY, SKIN									
PROC#: 11100 CPT-4: 11100 LINENO: 41									
REL VALUE	1.07	1.25	1.50	1.39	1.40	1.63	1.54	1.50	1.35
SCALE RANK	66.0	60.0	41.5	6.0	6.0	53.0	52.5	51.0	50.0
COMP OV EPAT									
PROC#: 90080 CPT-4: 90080 LINENO: 10									
REL VALUE	1.24	1.67	3.61	.	.	2.08	2.23	2.06	2.00
SCALE RANK	67.0	68.5	70.0	.	.	61.0	61.0	60.0	61.0
BRIEF HV NPAT									
PROC#: 90200 CPT-4: 90200 LINENO: 16									
REL VALUE	1.26	1.67	1.80	.	.	2.02	1.99	1.86	1.75
SCALE RANK	68.0	68.5	51.5	.	.	60.0	60.0	59.0	57.5

DIST POINT	HCFA MEAN 50TH	HCFA 75TH 50TH	MMA MMA	HIAA MEAN 50TH	HIAA 75TH 50TH	CAL78 MEAN BA	CAL78 75TH BA	CAL74 MEAN BA	CAL74 75TH BA
LTD CONSULTATION									
PROC#: 90600 CPT-4: 90600 LINENO: 26									
REL VALUE	1.26	1.67	1.80	.	.	1.86	1.92	1.67	1.50
SCALE RANK	69.0	68.5	51.5	.	.	58.0	58.5	54.0	53.0
SIGMOIDOSCOPY									
PROC#: 45300 CPT-4: 45300 LINENO: 55									
REL VALUE	1.33	1.67	1.75	1.36	1.24	1.57	1.54	1.49	1.31
SCALE RANK	70.0	68.5	47.0	5.0	5.0	52.0	52.5	50.0	48.0
COMP OV NPAT									
PROC#: 90020 CPT-4: 90020 LINENO: 4									
REL VALUE	1.42	2.00	4.81	.	.	2.66	2.88	2.60	2.50
SCALE RANK	71.0	71.5	74.0	.	.	68.0	69.0	67.0	68.5
SPINE X-RAY, COMPLETE									
PROC#: 72110 CPT-4: 72110 LINENO: 70									
REL VALUE	1.48	2.01	3.17	.	.	2.40	2.44	2.38	2.25
SCALE RANK	72.0	73.0	68.0	.	.	65.0	64.0	65.0	65.5
PSYCHOTHERAPY 50-60 MIN									
PROC#: 90803 CPT-4: 90844 LINENO: 29									
REL VALUE	1.56	2.00	.	.	.	2.18	2.23	2.27	2.00
SCALE RANK	73.0	71.5	.	.	.	63.0	62.0	62.0	61.0
THORACENTESIS									
PROC#: 32000 CPT-4: 32000 LINENO: 50									
REL VALUE	1.62	2.06	3.00	2.01	2.00	1.98	1.92	1.81	1.74
SCALE RANK	74.0	74.0	63.0	7.0	7.0	59.0	58.5	56.0	56.0
INTERMED HV NPAT									
PROC#: 90215 CPT-4: 90215 LINENO: 17									
REL VALUE	1.73	2.17	3.00	.	.	2.47	2.50	2.36	2.13
SCALE RANK	75.0	75.0	65.5	.	.	67.0	66.5	63.0	63.0
COMP HV NPAT									
PROC#: 90220 CPT-4: 90220 LINENO: 18									
REL VALUE	1.81	2.37	4.21	.	.	3.06	3.08	2.93	2.70
SCALE RANK	76.0	77.0	72.5	.	.	72.0	71.5	70.0	70.0

DIST POINT	HCFA		MMA	HIAA		CAL78		CAL74	
	MEAN 50TH	75TH 50TH	MMA	MEAN 50TH	75TH 50TH	MEAN BA	75TH BA	MEAN BA	75TH BA
EXTENSIVE CONSULTATION PROC#: 90610 CPT-4: 90610 LINENO: 27									
REL VALUE	1.85	2.33	3.00	.	.	2.44	2.50	2.37	2.18
SCALE RANK	77.0	76.0	65.5	.	.	66.0	66.5	64.0	64.0
UPPER GI TRACT X-RAY PROC#: 74242 CPT-4: 74240 LINENO: 72									
REL VALUE	1.95	2.40	3.70	.	.	2.91	2.90	2.80	2.50
SCALE RANK	78.0	78.5	71.0	.	.	71.0	70.0	69.0	68.5
COLON X-RAY, BARIUM ENEMA PROC#: 74270 CPT-4: 74270 LINENO: 73									
REL VALUE	1.97	2.50	2.12	.	.	2.74	2.72	2.57	2.35
SCALE RANK	79.0	80.5	58.0	.	.	69.0	68.0	66.0	67.0
CYSTOURETHROSCOPY PROC#: 52000 CPT-4: 52000 LINENO: 59									
REL VALUE	1.99	2.40	5.00	2.67	2.60	.	.	.	.
SCALE RANK	80.0	78.5	75.0	8.0	8.0	.	.	.	.
EEG PROC#: 95819 CPT-4: 95819 LINENO: 37									
REL VALUE	2.00	2.50	3.37	.	.	2.88	3.08	3.00	2.79
SCALE RANK	81.0	80.5	69.0	.	.	70.0	71.5	71.0	71.0
COMP CONSULTATION PROC#: 90620 CPT-4: 90620 LINENO: 28									
REL VALUE	2.31	2.83	4.21	.	.	3.40	3.46	3.22	3.00
SCALE RANK	82.0	82.0	72.5	.	.	73.0	73.0	72.0	72.0
BRONCHOSCOPY PROC#: 31620 CPT-4: 31620 LINENO: 49									
REL VALUE	6.52	7.46	9.00	8.76	8.01	10.38	9.71	9.01	7.80
SCALE RANK	83.0	83.0	76.0	9.0	9.0	74.0	74.0	73.0	73.0
CAT SCAN HEAD PROC#: 70470 CPT-4: 70470 LINENO: 77									

	HCFA		MMA	HIAA		CAL78		CAL74	
DIST POINT	MEAN 50TH	75TH 50TH	MMA	MEAN 50TH	75TH 50TH	MEAN BA	75TH BA	MEAN BA	75TH BA
REL VALUE	7.83	10.00	16.61	.	.	.	.	.	.
SCALE RANK	84.0	84.0	77.5	.	.	.	.	.	.
CAT SCAN ABDOMEN									
PROC#: 74170 CPT-4: 74170 LINENO: 78									
REL VALUE	8.80	13.10	16.61	.	.	.	.	.	.
SCALE RANK	85.0	85.0	77.5	.	.	.	.	.	.
HEMORRHOIDECTOMY									
PROC#: 46255 CPT-4: 46255 LINENO: 56									
REL VALUE	12.13	13.33	17.50	15.67	14.65	.	.	.	.
SCALE RANK	86.0	86.0	79.0	10.0	10.0	.	.	.	.
REPAIR HERNIA									
PROC#: 49505 CPT-4: 49505 LINENO: 58									
REL VALUE	14.17	16.50	22.50	17.63	16.03	23.06	23.01	20.82	18.20
SCALE RANK	87.0	87.0	80.0	11.0	11.5	75.0	75.0	74.0	74.0
APPENDECTOMY									
PROC#: 44950 CPT-4: 44950 LINENO: 54									
REL VALUE	14.41	16.63	25.00	18.41	16.03	24.50	23.89	21.72	19.31
SCALE RANK	88.0	88.0	82.0	12.0	11.5	76.0	76.0	75.0	75.0
CATHERIZATION, HEART, RIGHT&LE									
PROC#: 93527 CPT-4: 93527 LINENO: 51									
REL VALUE	16.09	20.61	24.04	.	.	.	.	.	.
SCALE RANK	89.0	89.0	81.0	.	.	.	.	.	.
COLECYSTECTOMY									
PROC#: 47600 CPT-4: 47600 LINENO: 57									
REL VALUE	21.71	25.37	41.50	28.30	25.44	36.33	34.62	33.13	29.06
SCALE RANK	90.0	90.0	84.0	13.0	13.0	77.0	77.0	76.0	76.0
EXTRACTION OF LENS (1)									
PROC#: 66920 CPT-4: 66920 LINENO: 67									
REL VALUE	27.29	30.00	50.00	35.36	32.05	43.82	42.31	41.52	37.09
SCALE RANK	91.0	91.0	87.5	15.0	15.0	78.0	78.0	79.0	79.0

DIST POINT	HCFA MEAN 50TH	HCFA 75TH 50TH	MMA	H1AA MEAN 50TH	H1AA 75TH 50TH	CAL78 MEAN BA	CAL78 75TH BA	CAL74 MEAN BA	CAL74 75TH BA
MODIFIED RADICAL MASTECTOMY									
PROC#: 19240 CPT-4: 19240 LINENO: 43									
REL VALUE	27.32	31.21	40.00	.	.	44.10	44.33	37.17	33.00
SCALE RANK	92.0	92.0	83.0	.	.	79.0	79.0	77.0	77.0
HYSTERECTOMY									
PROC#: 58265 CPT-4: 58265 LINENO: 63									
REL VALUE	27.98	33.33	52.50	34.26	32.05	46.12	46.15	39.55	35.02
SCALE RANK	93.0	95.0	90.0	14.0	15.0	80.0	80.0	78.0	78.0
PROSTATE, TRANSURETHRAL ELECTR									
PROC#: 52601 CPT-4: 52601 LINENO: 62									
REL VALUE	29.23	32.00	50.00	36.23	32.05	.	.	.	.
SCALE RANK	94.0	93.0	87.5	17.0	15.0	.	.	.	.
PROSTATECTOMY, SUPRAPUBIC									
PROC#: 55821 CPT-4: 55821 LINENO: 61									
REL VALUE	29.47	32.88	50.00	38.74	35.66	.	.	.	.
SCALE RANK	95.0	94.0	87.5	18.0	18.0	.	.	.	.
COLECTOMY, PARTIAL									
PROC#: 44140 CPT-4: 44140 LINENO: 53									
REL VALUE	31.62	36.45	45.00	39.50	36.06	52.69	50.77	45.61	40.00
SCALE RANK	96.0	97.0	85.0	19.0	19.0	81.0	81.0	80.0	80.0
FEMUR FRACTURE, OPEN REDUCTION									
PROC#: 27236 CPT-4: 27236 LINENO: 44									
REL VALUE	32.63	36.25	55.00	43.47	40.26	56.94	56.78	50.24	45.00
SCALE RANK	97.0	96.0	91.0	20.0	20.0	82.0	82.0	81.0	81.0
INSERTION OF PERMANENT PACEMAK									
PROC#: 33200 CPT-4: 33200 LINENO: 52									
REL VALUE	34.70	50.00	50.00	35.95	35.46	.	.	.	.
SCALE RANK	98.0	99.0	87.5	16.0	17.0	.	.	.	.
DONOR NEPHRECTOMY-LIVING DONOR									
PROC#: 50320 CPT-4: 50320 LINENO: 102									
REL VALUE	35.40	40.00	60.00	49.96	58.09	.	.	.	.

	HCFA		HMA	HIAA		CAL78		CAL74	
DIST POINT	MEAN 50TH	75TH 50TH	HMA	MEAN 50TH	75TH 50TH	MEAN BA	75TH BA	MEAN BA	75TH BA
SCALE RANK	99.0	98.0	92.0	21.0	21.0	.	.	.	.
ARTHROPLASTY PROC#: 27130 CPT-4: 27130 LINENO: 47									
REL VALUE	61.04	76.08	75.00	.	.	102.3	100	94.51	84.15
SCALE RANK	100	100	93.5	.	.	83.0	83.0	82.0	82.0
KIDNEY TRANSPLANT PROC#: 50360 CPT-4: 50360 LINENO: 103									
REL VALUE	63.21	78.49	75.00	.	.	.	.	.	.
SCALE RANK	101	101	93.5	.	.	.	.	.	.
CORONARY BYPASS >=3 ARTERIES PROC#: 33518 CPT-4: 33518 LINENO: 46									
REL VALUE	92.02	121.7	125	.	.	.	.	.	.
SCALE RANK	102	102	95.0	.	.	.	.	.	.